Author/s:
Andrejević, M;White, JP;Feuerriegel, D;Laham, S;Bode, S

Title:
Response time modelling reveals evidence for multiple, distinct sources of moral decision caution

Date:
2022-06-01

Citation:
Andrejević, M., White, J. P., Feuerriegel, D., Laham, S. & Bode, S. (2022). Response time modelling reveals evidence for multiple, distinct sources of moral decision caution. Cognition, 223, https://doi.org/10.1016/j.cognition.2022.105026.

Cognition 223 (2022) 105026

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Response time modelling reveals evidence for multiple, distinct sources of moral decision caution

Milan Andrejević [a,b,*], Joshua P. White [a], Daniel Feuerriegel [a], Simon Laham [a], Stefan Bode [a]

[a] Melbourne School of Psychological Sciences, The University of Melbourne, Australia
[b] Cognition & Philosophy Lab, Department of Philosophy, Monash University, Australia

ARTICLE INFO

Keywords:
Moral decision caution
Moral judgement
Response time
Drift diffusion model
Dictator game
Context update

ABSTRACT

People are often cautious in delivering moral judgements of others' behaviours, as falsely accusing others of wrongdoing can be costly for social relationships. Caution might further be present when making judgements in information-dynamic environments, as contextual updates can change our minds. This study investigated the processes with which moral valence and context expectancy drive caution in moral judgements. Across two experiments, participants ($N = 122$) made moral judgements of others' sharing actions. Prior to judging, participants were informed whether contextual information regarding the deservingness of the recipient would follow. We found that participants slowed their moral judgements when judging negatively valenced actions and when expecting contextual updates. Using a diffusion decision model framework, these changes were explained by shifts in drift rate and decision bias (valence) and boundary setting (context), respectively. These findings demonstrate how moral decision caution can be decomposed into distinct aspects of the unfolding decision process.

1. Introduction

Evaluations of people's actions or thoughts as morally good or bad (henceforth moral judgements) are an integral and pervasive part of being human – they undergird our social relationships and form the basis for our legal, political, and governmental institutions (Turiel, 2001). Yet moral judgements are not made in isolation but in a complex informational context; various types of contextual information can influence our moral judgements. For example, moral judgements can be modulated by contextual information regarding the social identities of moral actors and victims (Miron, Warner, & Branscombe, 2011; Sawaoka, Newheiser, & Dovidio, 2014), their economic class (Olson, McFerran, Morales, & Dahl, 2016), their relational status (Berg, Kitayama, & Kross, 2021; Earp, Mcloughlin, Monrad, Clark, & Crockett, 2020; Haidt & Baron, 1996; Simpson, Laham, & Fiske, 2016; Weidman, Sowden, Berg, & Kross, 2020), the actor's mitigating circumstances (Feather & Deverson, 2000), as well as the victim's history of moral or immoral actions (i.e., their moral deservingness; Andrejević, Feuerriegel, Turner, Laham, & Bode, 2020; Feather, 1999). Such information can also lead people to change their minds about their moral judgements. Recent research has shown that people flexibly update their judgements upon receiving contextual information, switching from relying on context-independent to context-dependent norms (Andrejević et al., 2020).

In distributive justice scenarios, context-dependent norms are often preferred over context-independent norms. For example, if information regarding individual contributions of actors to shared resources or the history of previous transactions is made available, people prefer splitting resources in accordance with norms that account for such information (e.g., equity norm; Deutsch, 1975; Hysom & Fişek, 2011; Shapiro, 1975; reciprocity norm; Diekmann, 2004; Fehr & Fischbacher, 2004; indirect reciprocity norm; Meristo & Surian, 2013; Nowak & Sigmund, 1998), rather than ignoring this information and allocating equal amounts to each individual (equality norm; Adams, 1965; Fehr & Schmidt, 1999; Messick & Schell, 1992). Some individuals prefer context-dependent norms so much that they refrain from making strong judgements prior to the presentation of contextual information (Andrejević et al., 2020). This may reflect the caution of these individuals not to make judgements that may later, upon learning additional information, turn out to be mistaken as they no longer align with preferred context-dependent norms. Thus, caution against selecting a judgement that is not in line with personal moral norms (i.e., moral decision caution) likely plays an

* Corresponding author at: Melbourne School of Psychological Sciences, The University of Melbourne, Australia.
E-mail address: milan.andrejevic@unimelb.edu.au (M. Andrejević).

https://doi.org/10.1016/j.cognition.2022.105026
Received 10 April 2021; Received in revised form 12 January 2022; Accepted 13 January 2022

important role in dynamic everyday moral decision-making situations, especially when we are aware that we may learn additional, decision-relevant information in the near future. However, the moral decision process in dynamic situations is poorly understood, partly because previous research has not employed process models to investigate various forms of moral decision caution more directly.

Caution has been studied in different areas of decision research, mostly using single-decision tasks involving perceptual and reward-based choices. One form of caution characterized by this research is a tendency for an individual to slow their response time (RT) in order to increase the likelihood of making a correct choice (Forstmann et al., 2010). Generally, participants show this tendency when under explicit instructions to ensure high accuracy (Voss, Rothermund, & Voss, 2004), when there are high monetary costs for mistakes (Green, Biele, & Heekeren, 2012), and in conditions of high task difficulty, so as to maximize reward rate (Starns & Ratcliff, 2012). This form of caution enables people to adapt their decision processes to suit changing environmental demands.

This form of caution may also be a useful way for people to adapt their moral decision-making when there is an expectancy of learning more information at the time of the decision. Recent research suggests that people display this form of caution as they learn about the likelihood of outcomes or consequences of their choices. People slow their judgements to reduce the likelihood of errors when they are aware that they are likely to make an error in the given task (Dunovan & Verstynen, 2019), as well as when they are aware that the association between their choices and the outcomes resulting from these choices is volatile (Bond, Dunovan, & Verstynen, 2018). Expectancy of learning contextual information in moral judgements also changes the subjective likelihood of judgement errors, defined as choices that appear correct based on the information available at the time of the decision (in line with context-independent norms) but turn out to be incorrect as contextual information is learned (not in line with the preferred context-dependent norms). Therefore, expectancy of learning contextual information may also lead to this form of caution and slow down moral judgements.

However, there is also another form of caution, which is highly relevant for third-person moral judgements of actions of others: People may particularly slow their RTs when judging someone else's action as morally bad, to increase the likelihood of being correct (according to their personal moral norms) when selecting this option. This tendency can be conceptualised as a *decision bias*, which has been shown to occur in other contexts against choice options associated with smaller rewards (Voss et al., 2004), or larger punishments (Summerfield & Koechlin, 2010). Morally blaming others is socially risky as it may lead to reprisals if that blame is improperly placed. Indeed, when making negative (as compared to positive) judgements, people are more motivated to stay as accurate as possible by ensuring their judgements are up to date with all the available information (Anderson, Crockett, & Pizarro, 2020; Monroe & Malle, 2019; Siegel, Mathys, Rutledge, & Crockett, 2018). However, there is an alternative explanation for why people may take longer when judging someone as bad that is unrelated to caution. Namely, people tend to take longer to evaluate negative information, even when they are not required to make any decisions, and there are no response options to be cautious about. For instance, people report thinking more thoroughly about negative events (Abele, 1985), they look longer at negative content when scrolling through images (Fiske, 1980), and are longer distracted by morally negative words (Wentura, Rothermund, & Bak, 2000). Such effects suggest that people take longer to process negatively valenced information (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Therefore, there are two distinct explanations for slower RTs when making third-party negative moral judgements: a decision bias (defined as a tendency to be more cautious when judging someone as bad), and a slower rate of evaluation of negative information (i.e., evidence for the negative judgement). For this reason, previous research relying on simple comparisons of mean RTs has been unable to disentangle the cognitive processes underlying the slowing (Baumeister et al., 2001). This is again due to the fact that there is no process model specifically developed for third-party moral judgements that would allow us to investigate this question.

In other fields of decision science, evidence accumulation models have been widely applied to disentangle parts of the decision process. These process models include mathematically formalized parameters that correspond to evidence accumulation (i.e., the rate at which evidence is evaluated) and the two forms of caution described above, and might therefore be useful for partitioning distinct sources of moral decision caution. One prominent model of this class is the Diffusion Decision Model (DDM; Ratcliff, 1978; Ratcliff, Smith, Brown, & McKoon, 2016; Smith & Ratcliff, 2004) which has been used to study decision-making across a broad range of discrete choice tasks (Gomez, Ratcliff, & Perea, 2007; Kiani & Shadlen, 2009; Ratcliff, 1978; Ratcliff, Gomez, & McKoon, 2004; Ratcliff, Thapar, & McKoon, 2004). The DDM describes the decision process as a continuous accumulation of noisy evidence for different choice outcomes. Once evidence in favour of a particular choice reaches a boundary, a decision is made. These models find substantial support from animal studies where neural firing rates in middle temporal and ventral intraparietal areas found to closely track the trajectory of evidence accumulation (Cook & Maunsell, 2002; Gold & Shadlen, 2007). Although predominantly used to model perceptual decision-making processes, where sensory evidence is accumulated by the sensory systems, the DDM can be regarded as a universal decision process model, and it has been used to model value-based decisions (Milosavljevic, Malmaud, Huth, Koch, & Rangel, 2010), sharing and cooperation choices (Gallotti & Grujić, 2019; Hutcherson, Bushong, & Rangel, 2015; Krajbich, Hare, Bartling, Morishima, & Fehr, 2015; Son, Bhandari, & Feldman Hall, 2019) as well as moral decisions (Pärnamets et al., 2015; Yu, Siegel, Clithero, & Crockett, 2021). Moreover, several authors have argued that RT distributions of moral choices in sacrificial moral dilemmas, when averaged across a large enough sample (aggregated across participants and studies), are generally consistent with predictions of the DDM framework (Baron & Gürçay, 2017; Cohen & Ahn, 2016). For such higher-level decisions, the accumulation process represents integration of signals from brain areas that calculate subjective value (Plassmann, O'Doherty, & Rangel, 2007), integrate representations of potential gains and losses (Basten, Biele, Heekeren, & Fiebach, 2010), and perform diverse social and moral computations that have not yet been well specified by previous research (Baron & Gürçay, 2017; Cohen & Ahn, 2016; Hutcherson et al., 2015; Pärnamets et al., 2015).

The rate of evidence processing (i.e., the evidence strength), and two forms of caution, correspond to specific parameters of the DDM model. In the DDM model, caution against making an error across response options is formalized as the amount of evidence needed to make a choice and is estimated by *the boundary separation (a)* parameter. Given the role of this parameter in adapting decision processes to environmental demands (Bond et al., 2018; Dunovan & Verstynen, 2019; Voss et al., 2004; as described above), this parameter may increase - corresponding to a wider boundary separation, when expecting contextual information. Biases against one of the response options are formalized as shifts in the starting point of the accumulation process, thus capturing how much people favour a certain response option prior to observing the stimulus, and is estimated by *the (starting point) bias parameter (z)*. This parameter has been shown to shift towards the response option associated with a reward (Voss et al., 2004) and away from response options associated with punishment (Summerfield & Koechlin, 2010). In the case of moral judgements, because of the potential social repercussions that come with placing moral blame improperly, the bias parameter may be shifted towards the "good" judgement choice. And third, the average rate of evidence accumulation, capturing the strength of evidence favouring either response option in a task, is estimated by *the drift rate (v)* parameter. In visual discrimination tasks, this parameter has been shown to scale with stimulus discriminability (Voss et al., 2004). We expected this parameter to scale with the prototypicality of action as morally good (representing adherence to a moral norm) or bad (representing deviation from a moral norm). Moreover, if negative evidence is accumulated slower than positive evidence (independent from potential biased caution against "bad" judgements accounted by the z parameter), we expect reduced absolute drift rates in trials with negatively valenced stimuli. This would support the idea that negative information is processed slower for reasons unrelated to caution per se.

In the current study we used a modified version of a recently developed third-party moral judgement task (Andrejević et al., 2020; Andrejević et al., 2021) to test the effects of context-expectancy and moral valence on RT and parameters of the decision process, as operationalised by the DDM. We asked participants to observe a variant of the dictator game in which a "Decision-maker" decided to share a proportion of $10 with another person (the "Receiver"; Fig. 1A). Participants were aware that these choices were made in a particular informational context. Participants knew either that decision-makers knew nothing about the Receiver, or that they knew how 'deserving' the Receiver was, based on their past sharing behaviour towards another person. In each trial participants made moral judgements about the Decision-maker's sharing action while expecting a contextual update about the Receiver's deservingness (*context-expectant* condition) or while not expecting an update (*no-context* condition; Fig. 1B). In Experiment 1 these two conditions were presented interleaved. Experiment 2 was used to replicate the results using a near-identical paradigm with an independent sample of participants. In the second experiment the two conditions were presented in separate blocks, which further controlled for the possibility that the interleaved presentation of conditions might have had an impact on participants' decision strategies. Naturally, there are individual differences in the norms people rely on to make such judgements. A majority of people, however, condemn low and endorse high offers (Andrejević et al., 2020). To avoid possible confounding of response times due to potential differences in the reliance on different sets of norms across individuals, and to ensure that the perception of our stimuli as "evidence" for judgement options was roughly consistent across the sample (a necessary assumption of the DDM when fit for a group of individuals in a hierarchical model, see Methods), we limited our investigation in both experiments to this largest subset of participants, who endorsed generosity and condemned selfishness (Andrejević et al., 2020; implications for limitations will be addressed below).

2. Method

2.1. Participants

The study was approved by the Human Research Ethics Committee of the Melbourne School of Psychological Sciences (Ethics ID 1750046.3). Participants were compensated with course credit or monetary remuneration ($15). Participants were right-handed, fluent in English, and had normal or corrected-to-normal vision.

For Experiment 1 (interleaved design), 77 people participated (50 female, 27 male, $M_{age} = 24.70$, $SD = 7.40$, range: 18–69 years). Eleven participants were excluded from the sample for data quality reasons: nine participants failed an attention-check (i.e., had given incorrect answers in more than 40% of catch trials of either category; see below), and two participants had missing responses for over 5% of trials, again suggesting a lack of attention. We preselected the final sample such that all included participants would rely on the same moral norms to make their judgements. This was done to avoid possible confounding of response times due to diversity of moral judgement strategies, and to ensure that all participants were assigning moral meaning to presented stimuli in a similar manner (which is a necessary assumption of the DDM when fit for a group of participant datasets). Based on previous research using a similar task, we expected the largest group to be participants who endorsed high and condemned low offers (Andrejević et al., 2020). A strong positive correlation between moral judgements and Decision-maker's offer (indicating that judgements scaled with the offer

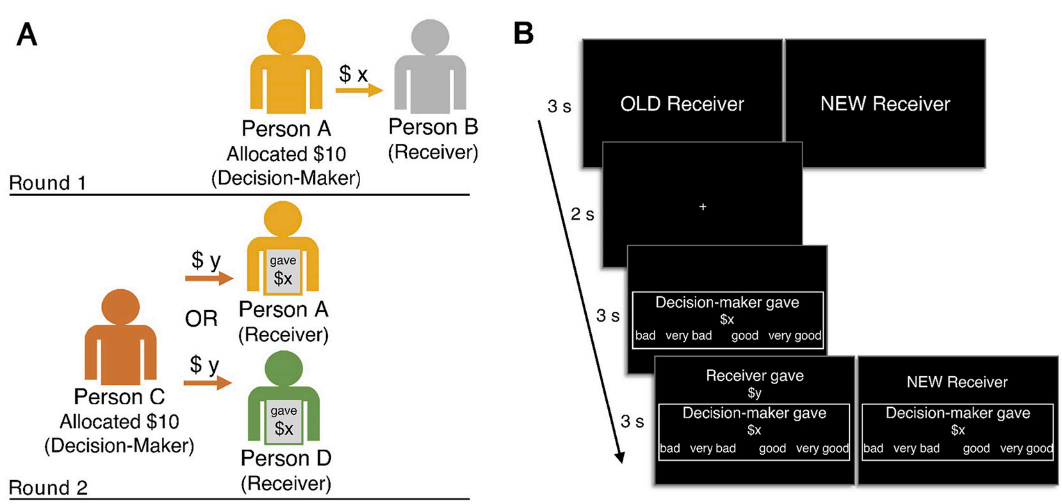

Fig. 1. Moral judgement context-expectancy paradigm.
Note. Panel A: Depiction of the cover story participants read prior to the experiment about a recently conducted study. The cover-story study was fictitious, but our participants were not informed of this. It involved persons interacting across two rounds: In Round 1, Person A played the role of the Decision-maker and had to decide how to share $10 with their partner, Person B. In Round 2, a new person (Person C) became a Decision-maker and was paired with either Person A or a new person (Person D), and had to decide how to share $10 with their partner. Importantly, Person C knew whether their partner took part in Round 1, and if they did (e.g., Person A), how much they gave when they were the Decision-maker (*$x*). Person C decided to give a certain amount (*$y*) to their partner (either person A or person D, depending on the trial). Panel B: Trial sequence. Participants were presented with information regarding the context-expectancy condition of the current trial. "OLD Receiver" indicated that they would judge the Round 2 Decision-maker who was paired with a Receiver (i.e., Person A) who gave an amount $x to another person in the previous round. Our participants made this judgement without yet knowing this $x amount, but knowing that they would soon learn this information (i.e., the context-expectancy condition); or "NEW Receiver", indicating that they would judge the Round 2 Decision-maker paired with a new person (i.e., Person D), and that there was no additional contextual information to expect (no-context condition). Next, participants were presented with the amount that the Round 2 Decision-maker gave to their partner and selected their judgement (one of the four options) on a keyboard. After this, in context-expectant condition, the amount that the Receiver had given in the previous round (*$x*) was revealed. In the no-context condition, no additional information was presented. Participants again indicated their judgement of the Decision-maker's action on their keyboard.

magnitude) was typical for this largest group. We excluded eleven participants who did not show this strong positive correlation (Spearman correlation was below $r = 0.5$). We report patterns of responses of participants excluded based on this criterion in Supplement 4. All of these criteria were predefined and preregistered (http://aspredicted.org/blind.php?x=n2fi7g). The final sample consisted of 55 participants (37 female, 18 male, $M_{age} = 24.84$, $SD = 5.86$, range, 18–43 years).

For Experiment 2 (blocked design), 76 members of the University of Melbourne community were recruited (47 female, 28 male, 1 other, $M_{age} = 24.29$, $SD = 3.77$, range: 19–39 years). Nine participants were excluded to ensure data quality: six participants failed the attention-check criterion (see above) and three had missing responses for over 5% of trials. Another ten participants were excluded because their moral judgements did not correlate strongly with the Decision-maker's offer (Spearman correlation $r < 0.6$). All of these criteria were predefined and preregistered (https://aspredicted.org/blind.php?x=dy3qk9). The final sample consisted of 57 participants (38 female, 18 male, 1 other, $M_{age} = 24.34$, $SD = 3.80$, range: 20–39 years).

2.1.1. Apparatus

The experimental task was programmed in MATLAB (MathWorks, version R2015b) and presented using PsychToolbox-3 (Kleiner et al., 2007). Participants sat at a viewing distance of approximately 80 cm from the monitor (ASUS ROG Swift PQ258Q 24.5" HD with a 60 Hz screen refresh rate). The experiment was conducted in a well-lit solitary room. Participants made responses on a black Hewlett-Packard KU1469 QWERTY keyboard. The “z”, “x”, “.” and “/” keys were covered with white stickers to indicate to participants that these were the primary buttons to be used in the experiment. They were instructed to place their fingers on these keys in preparation for every trial in the following manner: the middle finger and the index finger of their left hand were to be placed on the “z” and “x” keys, respectively, and the index finger and the middle finger of their right hand were to be placed on the “.” and “/” keys, respectively.

2.2. Experimental paradigm

2.2.1. Cover story

Participants first read a cover story about a recently conducted experiment investigating people's economic decisions. This experiment was fictional, but participants were not informed of this. In the fictional experiment a group of people, assigned to pairs, completed a two-round variant of the dictator game (for the original dictator game, see Güth, Schmittberger, & Schwarze, 1982). In the first round, one person (the “Decision-maker”) in each pair was given \$10 and decided how much thereof to share with their partner, the “Receiver”, to whom they could give any whole dollar portion (i.e., any amount \$0–\$10). In the second round, the same task was repeated except with people taking new roles — first round Decision-makers became Receivers in the second round — and were assigned different partners. Some of these new partners were Decision-makers in the first round of the experiment (“Old Receivers”) and some of them were not (“New Receivers”). Importantly, second round Decision-makers were aware whether their partner was an Old Receiver or a New Receiver. If their partner was an Old Receiver, they were also aware how much money their partner had shared with another person in the first round of the experiment. A visualisation of this cover story is shown in Fig. 1A.

2.2.2. Instructions

This cover story along with the description of the experimental task were presented to participants via text interleaved with animated depictions. Participants read the instructions and attended to animations at their own pace. Participants were then required to pass, with 100% accuracy, a test comprised of 32 true–false questions which assessed their understanding of both the cover story and the experiment instructions. Participants could attempt this instruction-check test three times. If they experienced troubles completing the quiz, participants could return to the cover story or instruction presentations to clarify their understanding or ask questions of the experimenters for the same. Participants were required to pass this test before continuing to the experiment.

2.2.3. Experimental task

Participants were asked to observe a series of independent transactions that various Decision-makers made towards various Receivers as described in the cover story. Each trial started with the participant being shown, for 3 s, whether the Receiver for that trial was an “OLD Receiver” (for context-expectant trials) or a “NEW Receiver” (for no-context trials) which corresponded to whether the Receiver participated in the first round of the fictitious experiment. Then, a fixation cross was presented in the middle of the screen for 2 s. Participants were then presented with the phrase “Decision-maker gave: \$y” where y was an integer from the set $Y = \{0, 1, 2, \dots 10\}$ (“Decision-maker offer”). Simultaneously, response options “very bad”, “bad”, “good” and “very good” were presented below the Decision-maker offer. Participants selected their response, with a maximum response window of 3 s, to indicate how morally good or bad they believed this Decision-maker's action was by pressing the button on the keyboard corresponding to the position of the presented option. To control for possible RT differences that could arise due to differences in motor execution across different fingers, participants were randomly assigned one of four possible mappings of responses to buttons, and this mapping remained the same throughout the experiment. Four mappings were selected to ensure that across participants any of the four fingers was mapped onto each response option. For consistency, none of the mappings had a monotonically increasing or decreasing order in space.

Once participants made their response, the corresponding response option immediately changed colour to yellow until the 3 s time-limit had elapsed (or for 0.3 s if the judgement was made between 2.7 s and 3 s) before reverting to white. This was done to assure participants that their response had been recorded.

Participants were then shown another fixation cross above this information for 0.5 s. The stimuli presented next differed depending on the experimental condition of the trial. In context-expectant trials, participants were presented with the phrase “Receiver gave: \$x”, where x was an integer from the set $X = \{0, 1, 2, \dots 10\}$, providing the contextual information of how much the Receiver had given when they were a Decision-maker the first round. In the no-context trials, participants were presented with the phrase “NEW Receiver”, reminding them that the Receiver had not participated in the first round, and thus there was no contextual information about them available. In both conditions, participants made a second moral judgement, within 3 s, about the Decision-maker's action (not the Receiver's prior action). Once this response had been made, the corresponding response option changed colour to yellow until the 3 s time-limit had elapsed (or for 0.3 s if the judgement was made between 2.7 s and 3 s), after which a new trial began. Please note that data relating to the second judgement in each trial was not the focus of the current publication but will instead be analysed for a separate publication in the future.

There were 121 trials in each condition, totalling 242 trials per participant. This was chosen such that in the context-expectant condition, participants made moral judgements about all possible combinations of the Decision-maker's offer (i.e., Decision-maker gave \$0–\$10) and the Receiver's prior offer (Receiver gave \$0–\$10; $11 \times 11 = 121$). To ensure there was symmetry between the experimental conditions, we also included 121 trials for the no-context condition. In Experiment 1 the order of these 242 trials was randomised for each participant and the two trial types alternated randomly (i.e., the two conditions were interleaved). In Experiment 2, we used a version of the experiment with the two expectancy conditions presented in separate blocks. There were 40–41 trials of the same kind in each block and 6 alternating blocks in total. The order of trials was randomised for each participant, and the participants were randomly assigned one of the two alternating block sequences.

2.2.3.1. Questionnaires. Following the experiment, participants completed various personality measures. We administered the agreeableness section of the HEXACO Personality Inventory-Revised (HEXACO; Lee & Ashton, 2018), a brief set of self-report measures for political orientation (Graham, Haidt, & Nosek, 2009), the Social Dominance Orientation scale (SDO; Ho et al., 2015), the Consequentialist Thinking Scale (CTS; Piazza & Sousa, 2013), and basic demographic measures. We will analyse and report the questionnaire results in a separate publication.

2.2.3.2. Experiment feedback and instruction checks. Participants were instructed to respond as quickly and accurately as possible and always give a response. If they failed to do so within the 3 s time limit, they were presented with feedback at the end of that trial advising which response was missing (or both) and to "please make sure you always respond". Two types of attention-checks were also dispersed throughout the experiment. In one, participants were required to report the values seen in the current trial; that is, the amounts that the Decision-maker and/or the Receiver had given. Participants responded by entering this value into number keys on the keyboard. For the second attention-check participants had to report, via button press, whether the Receiver in the current trial was an Old Receiver or New Receiver. Participants were instructed that both these attention-check trials would occur at random times during the experiment.

2.3. Statistical analyses

2.3.1. Regression analysis

RTs for the first moral judgement were modelled with the Generalised Linear Mixed Models (GLMMs) approach which is a form regression suitable for hierarchical data (e.g., data of multiple individuals in several conditions) that is not normally distributed. Invalid trials (i.e., trials without any response) were excluded from all the analyses (0.72% of all trials in Experiment 1, and 1.17% in Experiment 2). GLMMs are superior to the common practice of transforming data before applying an ordinary-least-square linear mixed model (Lo & Andrews, 2015). GLMMs were specified as follows: An identity link was used because it assumes that RTs are direct measures of the duration of the decision process, rather than functional transformations of this duration (Lo & Andrews, 2015). A gamma distribution was used as the conditional distribution as it provided a good empirical fit to the data. Moreover, gamma distributed GLMMs have been used in numerous RT studies with similar tasks (De Boeck & Jeon, 2019; Maris, 1993; Palmer, Horowitz, Torralba, & Wolfe, 2011; Van Zandt, 2002). Lastly, random effects were included in the model to account for individual differences.

We compared a list of theoretically plausible candidate models which were derived with an increasingly complex random effects structure, as shown in Supplement 1 Table S1. For each random effect structure, a model was fit both with and without a fixed interaction parameter. For all models, the random effects were allowed to correlate; that is, the model had an unstructured variance-covariance matrix. Model parameters were estimated using maximum likelihood estimation via the Laplace approximation, implemented with the *glmmTMB* package (Brooks et al., 2017) in the R statistical programming environment (version 3.6.1). We selected the best fitting model using the Akaike Information Criterion (AIC). AIC was preferred over the likelihood ratio test, because not all compared models were nested, and because, unlike the likelihood ratio test, the AIC method helps prevent overfitting (Akaike, 1974). AIC was also preferred over the Bayesian Information Criterion (Schwarz, 1978) because it was unlikely that any of our candidate models are the true model, which better agreed with the assumptions of AIC (Vrieze, 2012). Akaike weights (Wagenmakers & Farrell, 2004) were calculated for all candidate models as a means to quantify the relative merits of the competing models, and the degree to which one model should be preferred over the others. Confidence intervals (and where necessary, p values) for fixed effects were calculated for most models using Wald's z method (Bolker et al., 2009). The fixed parameter effects from the best fitting model, and their 95% confidence intervals, were then used to test our hypotheses.

2.3.2. Diffusion Decision Model fitting

Participants' RT and decision data were fit in the Python 3.6 programming environment on a High-Performance Computing Cluster (Meade, Lafayette, Sauter, & Tosello, 2017), using the Hierarchical Drift Diffusion Model (HDDM) package (Wiecki, Sofer, & Frank, 2013). We used the HDDM.hddm function, which implements a hierarchical Bayesian Markov Chain Monte Carlo (MCMC) estimation of the DDM with four free parameters (a, v, z, and t). HDDM estimates these parameters for each individual, as well as at the group level (which are the estimates we report in this publication). This analysis was not preregistered but was run separately for Experiment 1 and Experiment 2 samples allowing us to assess whether the findings replicated across samples. Estimation procedure implemented in the HDDM package was chosen as it outperforms other estimation techniques and can accurately recover model parameters based on a small number of observations per participant, especially for participant sample sizes larger than 20 (Wiecki et al., 2013). Since the DDM is sensitive to outliers, it is recommended to devise exclusion criteria that ensure that some of the contaminant RTs are excluded whilst ensuring that criteria do not exclude larger portions of the data (e.g., more than 1% Ratcliff & Tuerlinckx, 2002). In line with recommendations from Ratcliff and Tuerlinckx (2002), we excluded trials in which reaction time was faster than 0.2 s (0.05% of valid trials in Experiment 1 and 0.27% of valid trials in Experiment 2), and slower than 2.8 s (0.37% of valid trials in Experiment 1 and 0.37% of valid trials in Experiment 2). Note, however, that including trials with reaction times slower than 2.8 s did not change the results (data not shown). The DDM was designed for binary decisions (e.g., "good" versus "bad"), which means that in order to model our data using the DDM, we simplified our data by collapsing across "very good" and "good" responses (*good* judgement) and across "very bad" and "bad" responses (*bad* judgement). We formulated two models to address our hypotheses: *m0* – the null model which assumes no difference between conditions when estimating DDM parameters; *m1* – the hypothesised model, in which parameter a was allowed to vary across two expectancy conditions ($a_{context\text{-}expectant}$ and $a_{no\text{-}context}$), and parameter v was fit to each discrete magnitude of Decision-Maker's offers separately ($v_{0\text{–}10}$, with 0 to 10 as nominal, discrete values). To estimate these parameters across conditions we used the '*depend_on*' argument within the HDDM.hddm function. We also report an alternative way of performing these comparisons of parameters across discrete conditions using the HDDM. Regression function which yielded the same results (see Supplement 5). For our Bayesian parameter estimation, we used the default non-informative priors in the HDDM package (Wiecki et al., 2013). This is the recommended option for novel tasks that are substantially different from typical perceptual decision-making paradigms prominent in the DDM literature (Wiecki et al., 2013). We obtained parameter estimates by generating a chain of 2500 MCMC samples of the joint Bayesian probability posterior distributions of all parameters at both participant and population level, and discarding the first 500 samples (as recommended in Wiecki et al., 2013). We evaluated chain convergence using Gelman-Rubin diagnostic over five repeated chains ($\hat{R}<1.1$ for all parameters and at all levels across Experiment 1 and Experiment 2). The two models – our theoretically plausible *m1* and the null model *m0* – were compared using the Deviance Information Criterion (DIC) goodness of fit measure, which penalises for model complexity. Additionally, we also assessed goodness of fit by performing the posterior predictive check procedure, by which we generated simulated data based on posteriors estimates and compared it to empirically observed data

(Supplement 2 Figs. S8–10). After establishing that the m1 model outperformed the null model and provided an excellent fit for our data, we tested our specific hypothesis regarding *a, v* and *z* parameters by directly comparing the Bayesian probability posteriors generated by the above-described MCMC procedure. We additionally tested whether context-expectancy influences effects of Decision-maker offer on observed drift rates in a post-hoc analysis (detailed in the Supplement 3).

2.4. Data availability

Data of all participants, materials including the instructions and the task code, as well as the analyses scripts that support the findings of this study are publicly available on an Open Science Framework (OSF) repository (DOI:10.17605/OSF.IO/EPD63; Andrejević, White, Feuerriegel, Laham, & Bode, 2021).

3. Results

The selected sample of participants relied on similar norms when performing their judgements (judging low offers as bad and high offers as good) as expected, and there were no systemic differences in the proportions of moral choice for each choice option across expectancy conditions (depicted in Fig. 2).

We took two approaches to test for effects of context-expectancy and moral valence on response speed in moral judgements. The first approach was to test for these effects by comparing RTs without formally specifying the decision process. The second approach was to use the DDM to better characterise these effects by comparing model parameter estimates across expectancy and valence conditions.

With regards to our first approach, we tested three hypotheses. First, we investigated whether expectancy of contextual information increases caution, by testing whether the RTs of initial judgements were higher in the context-expectant than in the no-context condition. Second, we investigated whether morally negative evidence is evaluated more cautiously and is processed at a slower rate. This hypothesis was operationalised as the assumption that the effect of morally negative valence linearly decreases with the size of the Decision-maker's offer. We therefore expected a negative relationship between the Decision-maker's offer and RT. Third, to investigate whether caution when expecting a contextual update is particularly pronounced for negative judgements, we tested whether the slope of the negative relationship between RT and Decision-maker's offer was steeper in the context-expectant condition. To test these hypotheses, we compared several GLMMs, which included the Decision-maker's offer, the expectancy condition, and their interaction, as predictors of RT (Supplement 1 Table S2).

The best-fitting model included main effects of Decision-maker's offer and expectancy condition but did not include an interaction, and the intercepts and the slopes of these two main effects were allowed to vary across individuals (Supplement 1 Table S2). The two main effects were substantial and statistically significant. Context-expectancy led to a 23 ms slowing of RTs, 95% CI [5, 41] in Experiment 1. One possibility is that this effect may have been reduced due to the interleaved design, which could have led to an overspill of decision criteria among trials of different conditions. This was addressed with Experiment 2, which used a blocked design and replicated the context-expectancy effect, which was indeed much larger. Context-expectancy led to a 138 ms slowing, 95% CI [110, 165]. These results support the hypothesis that context-expectancy increases caution. Regarding the second hypothesis, we found that a single dollar reduction in the Decision-maker's offer predicted a 26 ms slowing of RTs, 95% CI [31,21] in Experiment 1. This effect again replicated in Experiment 2 showing 28 ms slowing for each dollar reduction in the Decision-maker's offer, 95% CI [31, 24]. These results support the hypothesis that negatively valenced evidence is evaluated more cautiously or takes longer to process. The two main effects were consistent across quantiles (RT quantiles by condition are displayed in Fig. 3). They were also robust across different models (Supplement 1 Table S2) and across alternative approaches to modelling RT distributions (Supplement 1 Table S4 and S5). As for the interaction effect, there was no evidence across these two studies supporting the hypothesis that effects of negative valence are more pronounced when people are expecting a contextual update. Models including the interaction effect had poorer fits, and confidence intervals for this interaction parameter consistently included zero. In Supplement 3, we present additional analyses showing that effects of negative valence persist after we control for judgement extremeness (with "very bad" and "very good" considered as extreme options, and "bad" and "good" considered as moderate options). Given the possibility of quantitative differences in the relationship between offer magnitude and moral judgements for low ($0–4) and high ($6–10) offers (see Fig. 2), we additionally conducted a post-hoc regression which showed no evidence for a difference in slopes across the two offer ranges (Supplement 3).

Next, to better characterise these patterns of RT effects, and to test our predictions regarding the relationships between context-expectancy, moral valence, and components of the decision process, we fitted a DDM. To test whether context-expectancy increased the general amount of caution across judgement options (i.e., boundary separation), we computed two *a* parameters, one for each expectancy condition, and compared them. We expected: $a_{context\text{-}expectant} > a_{no\text{-}context}$. To test whether moral prototypicality of Decision-maker's offers reflected stronger evidence for judgement options (with lower offer magnitude reflecting evidence for "bad" option, and higher offers reflecting stronger evidence for the "good" option, in line with the norms applied by the selected sample), we fitted a *v* parameter separately for each Decision-maker's offer. The *v* parameter was signed, meaning negative values indicated evidence for "bad" judgement and positive values indicated evidence for "good" judgement. We tested for a monotonic positive relationship between the offer magnitude and the *v* parameter. Moreover, to test whether negatively valenced evidence is accumulated more

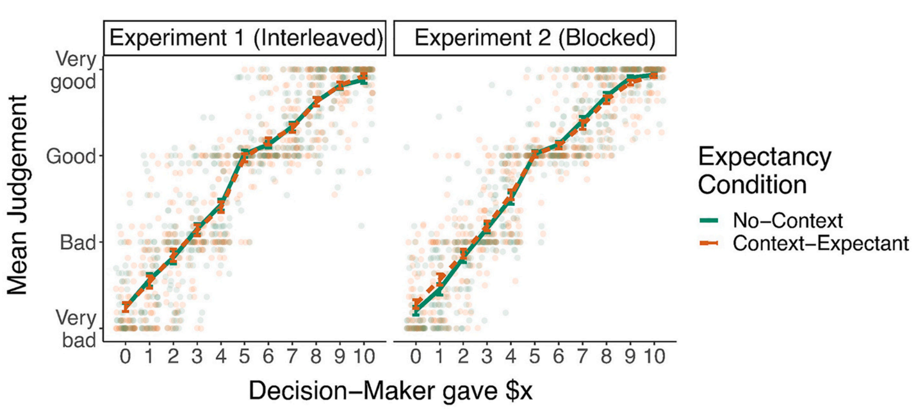

Fig. 2. Moral judgement results.
Note. Scattered dots indicate mean judgements of Decision-maker's offer for each participant; lines indicate the mean across participants for context-expectancy (orange) and no-context (green) conditions. Error bars depict the SEM. The monotonic increase in moral endorsement across Decision-maker's offers indicates that participants condemned low and endorsed high offers. The complete overlap of two the lines representing the conditions indicates that there were no detectable systemic differences in moral judgement across expectancy conditions in the two experiments.

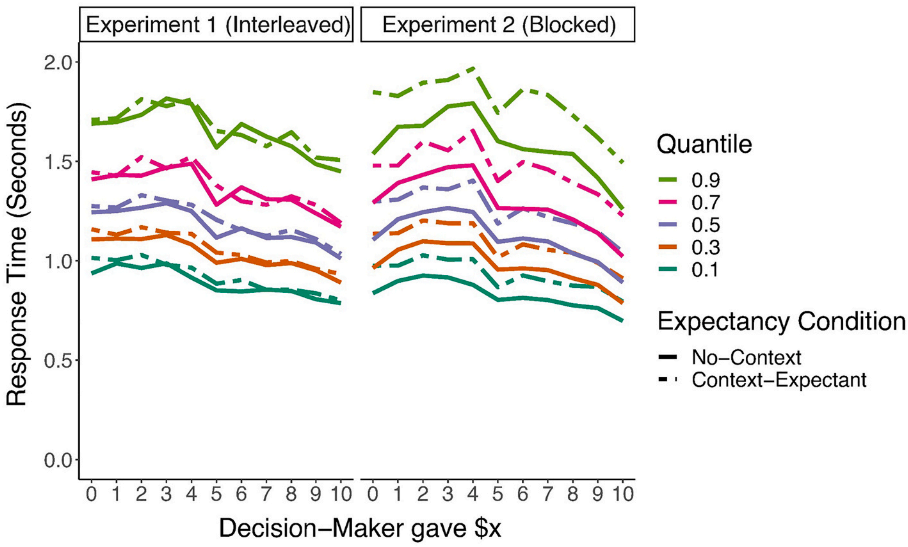

Fig. 3. Moral judgement RT quantiles across decision-maker offers and expectancy conditions.
Note. The graph shows group mean quantile values across participants. The general pattern of results was consistent across quantiles and across two studies: there was a slight increase in speed for higher Decision-maker offers; and there was a slight slowing in context-expectancy trials in Experiment 1 (dashed lines higher than solid lines), which was more pronounced in Experiment 2.

slowly than positively valenced evidence (independent from potential biased caution against "bad" judgements accounted by the z parameter), we tested whether the estimates of the v parameter were in absolute terms (drift towards either "good" or "bad") larger for high as opposed to low Decision-maker's offers. We expected: $|v_{0\text{–}4}| < |v_{6\text{–}10}|$. Finally, to test whether participants were more cautious against making "bad" judgements, independent of the tendency to more slowly accumulate negatively valenced information, we tested whether the z parameter differed from 0.5 (which would indicate no starting point bias), and whether the z parameter was biased in the direction of 'morally good' judgement. The position of decision bounds with respect to the starting point were standardized as 1 for 'morally good', and 0 for 'morally bad' judgements, hence we expected $z > 0.5$.

First, we formulated a hypothesised model (*m1*), which included separate a parameters for each expectancy condition, separate v parameters for every offer value, and a z parameter. We then tested whether the use of this model, which allowed us to test our specific hypotheses, was justifiable and appropriately explained our data, by comparing it to a null model (*m0*), which did not include differences between conditions for any parameter. We used the DIC to compare the model fits (lower value indicates better fit; Spiegelhalter, Best, Carlin, & Van Der Linde, 2002). We found that this model provided a substantially better fit to the data (Experiment 1 DIC = 13,396.303; Experiment 2 DIC = 15,021.171) than the null model (*m0*; Experiment 1 DIC = 29,569.801; Experiment 2 DIC = 30,071.387). Additionally, we ran Posterior Predictive Check (PPC) simulations of the two models. Simulated data from model *m1* more closely resembled the quantile structure of the observed RT data (Supplement 2 Fig. S10). The *m1* model simulation also reproduced the observed rates of judgements across Decision-maker offers (Supplement 2 Fig. S8) and patterns of changes in RT distributions across different Decision-maker's offers and expectancy conditions (Supplement 2 Fig. S9), and overall provided an excellent fit to the data.

Next, we tested for hypothesised differences in the *m1* model parameters across conditions. Consistent with our hypothesis that context-expectancy increases caution against making errors, a direct comparison of the posteriors yielded greater than 91% confidence that the a parameter estimate was larger in the context-expectant condition compared to the no-context condition in Experiment 1: posterior $P(a_{context\text{-}expectant} > a_{no\text{-}context}) = 0.913$ (Fig. 4A). In Experiment 2 confidence was greater than 99%: posterior $P(a_{context\text{-}expectant} > a_{no\text{-}context}) > 0.999$ (Fig. 4B). As for the drift rate (v), we expected this parameter to monotonically increase with the value of the Decision-maker's offer. We observed a perfect monotonic relationship across both experiments (see Fig. 4C and D). Next, we tested the hypothesis that there should be a reduction in absolute drift rate when processing negative moral valence as compared to positive valence. To do this, we computed a mean of v posterior parameter estimates across negatively valenced stimuli (i.e., Decision-maker gave \$0–4), and another mean of v parameter estimates across positively valenced stimuli (i.e., Decision-maker gave \$6–10), for each sample in the MCMC chain. The resulting posterior distributions for positive stimuli and for negative stimuli were then compared directly to test the hypothesis (as recommended in Wiecki et al., 2013). Consistent with our hypothesis, the comparison yielded a greater than 99% confidence that there was a decrease in absolute drift-rate for negative stimuli: Experiment 1 posterior $P(|v_{0\text{–}4}| < |v_{6\text{–}10}|) > 0.999$; Experiment 2 posterior $P(|v_{0\text{–}4}| < |v_{6\text{–}10}|) > 0.999$ (see Fig. 4C and D). To ensure that this effect was not due to a perception of Decision-maker's offers of \$4, and \$3 as neutral as opposed to negative, we repeated these analyses on a more constrained set of stimuli by excluding offers \$3–7, and the effect survived in both studies: Experiment 1 posterior $P(|v_{0\text{–}2}| < |v_{8\text{–}10}|) > 0.999$; Experiment 1 posterior $P(|v_{0\text{–}2}| < |v_{8\text{–}10}|) > 0.999$. To test our hypothesis regarding the shift of the bias parameter (z) away from the 'bad' and towards the 'good' judgement option, we tested whether the z parameter was larger than 0.5. Consistent with our hypothesis we found estimates of z parameter to be larger than 0.5 in both studies with confidence greater than 99%: Experiment 1 posterior $P(z > 0.5) > 0.999$; Experiment 2 posterior $P(z > 0.5) > 0.999$.

4. Discussion

We investigated the effects of context-expectancy and negative moral valence on moral decision caution in third-party moral judgements of sharing actions. Both factors were hypothesised to slow moral judgements, albeit impacting different aspects of the decision-making process. Specifically, we examined these effects by comparing RTs and parameter values of the DDM across judgements of fairness-related actions (i.e., offers of different magnitudes) as well as context-expectancy conditions. Our results show a significant slowing of RT in the context-expectancy conditions, as well as for morally negative actions; however, there was no interaction between the two factors. Moreover, these effects were well accounted for by differences in multiple DDM parameters. The boundary separation parameter was larger in the context-expectancy condition compared to the no-context condition, pointing to more caution to avoid erroneous responses (across judgement options) in the former condition. In addition, signed drift rates increased with the Decision-maker's offer, suggesting that lower offers corresponded to stronger evidence for negative judgements and higher offers

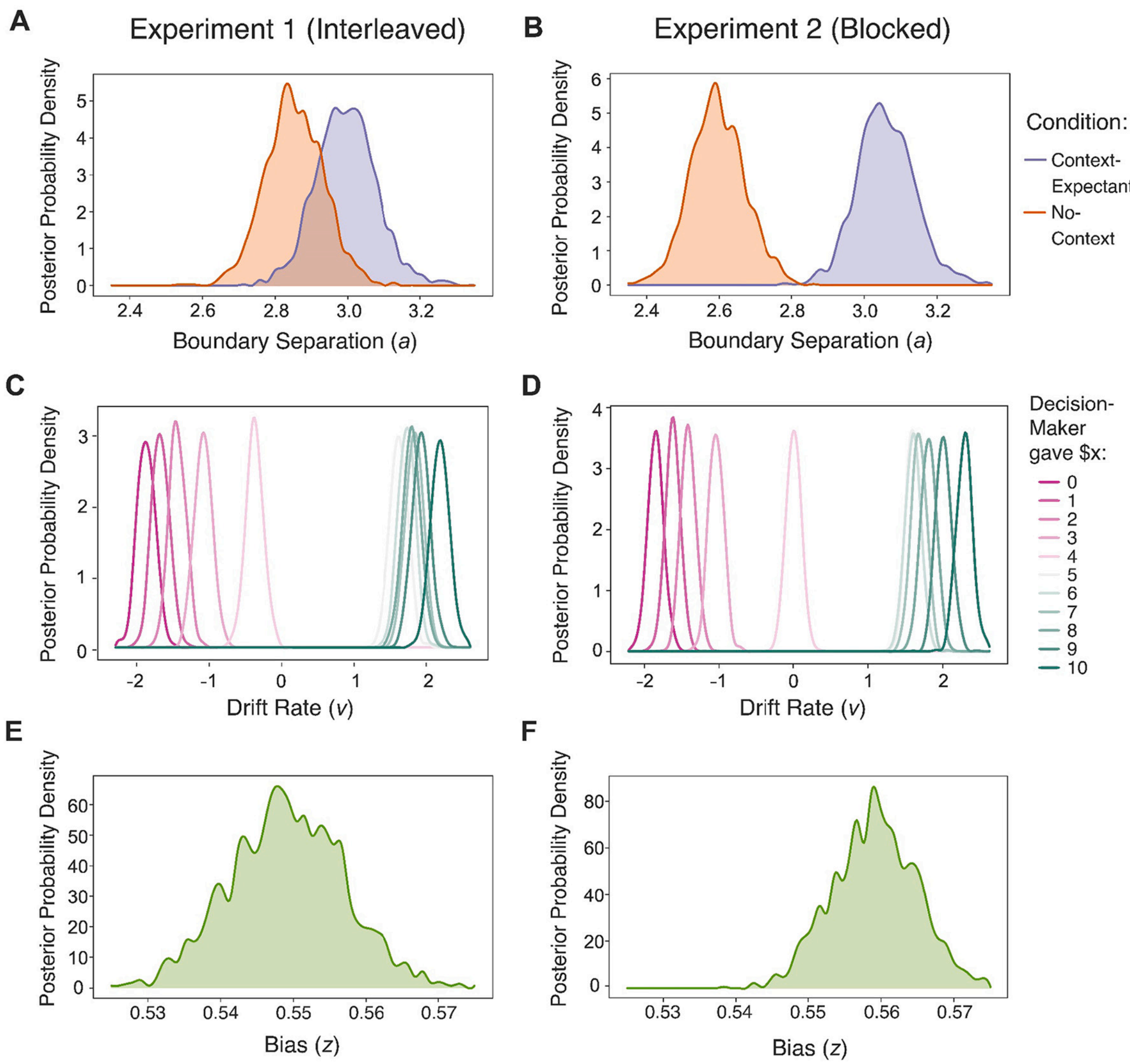

Fig. 4. Bayesian posterior probability distributions for diffusion decision model parameters a, v, and z for both experiments.
Note. Panel A: In Experiment 1, the boundary separation parameter (*a*) estimate, although overlapping, was slightly higher for the context-expectant condition, which is in line with the hypothesis that context-expectancy increases caution. Panel B: In Experiment 2, this difference was replicated with a larger effect and there was minimal overlap between the two posterior distributions. Panel C: In Experiment 1, the drift rate parameter (*v*) monotonically increased with higher Decision-maker's offers, suggesting that higher offer numbers provide more evidence for the judgement option 'good' and less for 'bad'. Positively valenced actions (DM gave more than 6) had higher absolute drift rates towards option 'good' than negatively valenced actions did towards option 'bad' (DM gave less than $4), which suggests that participants processed negatively valenced actions slower than positively valenced actions. Panel D: These effects replicated in Experiment 2. Panel E: In Experiment 1, participants showed a bias towards judging 'good' (z parameter >0.5), which is in line with the hypothesis that people may be more cautious when making negative judgements. Panel F: This effect replicated in Experiment 2.

corresponded to stronger evidence for positive judgements. Absolute drift rates were smaller for negatively valenced offers, supporting the notion that negative evidence is accumulated at a slower rate than positive evidence, for reasons most likely not related to caution. Additionally, the starting point parameter showed a bias against "bad" judgements, suggesting that people also slowed their negative judgements as they were particularly cautious about them.

Our findings that participants slowed their judgements when expecting contextual information is consistent with previous research showing that people are more cautious when aware that they are more prone to making mistakes (Bond et al., 2018; Dunovan & Verstynen, 2019). Notably, previous research has demonstrated this effect for decision mistakes in tasks in which people are not given additional information or a chance to change their minds (Bond et al., 2018; Dunovan & Verstynen, 2019). The current findings show that this effect also extends to dynamic decision-making contexts, in which learning additional information can lead to changes of mind. Crucially, here we show that this type of caution can be explained by the widening of the decision boundary separation in a process model of decision-making.

Finding that the expectancy of contextual information increases the boundary separation also highlights the importance of contextual information for moral judgements. This finding is consistent with previous research that showed that contextual information influences the judgements that we make (Feather, 1999; Feather & Deverson, 2000; Haidt & Baron, 1996; Miron et al., 2011; Olson et al., 2016; Sawaoka et al., 2014; Simpson et al., 2016), and that some people make less extreme good/bad judgements when expecting contextual information (Andrejević et al., 2020). Our confidence in this effect was smaller in the interleaved design, which may be because the effect of caution was weaker. Switching between expectancy conditions may be a demanding task that results in switching errors (Ethridge, Brahmbhatt, Gao, McDowell, & Clementz, 2009; Stoet & Snyder, 2007). Moreover, participants may adjust their decision-making to the decision task over several trails in the blocked condition, while interleaved design does not allow participants. To note, we did not find an adjustment of the judgement itself (see Fig. 2), but the relatively course four-point scale might not have been ideal to capture any potential subtle effects that might have occurred but could not be expressed without a finer scale. The difference in response times, however, was observed even though the expected contextual information could never directly impact the initial judgement. This is important because it shows that context-dependent norms affect our judgements even when contextual information is not yet known, a point which has been overlooked in the moral judgement literature.

We further found that participants were slower when evaluating lower offers, which is in line with both the idea that people take longer to process negative evidence (Abele, 1985; Baumeister et al., 2001; Fiske, 1980; Wentura et al., 2000), as well as with the idea people are more cautious against judging people as bad, as negative judgements have higher social repercussions for individuals (Monroe & Malle, 2019; Siegel et al., 2018). Our DDM results further support each of these accounts separately. Firstly, our finding that the drift rate was slower for lower offers as compared to higher offers is in line with the idea that people accumulate negative evidence at a slower rate (Abele, 1985; Baumeister et al., 2001; Fiske, 1980; Wentura et al., 2000). Secondly, we found that participants showed biases, or caution, against judging moral actions as bad, independent of taking longer to process negative evidence. Previous research on financial decision-making showed similar bias parameter shifts away from options associated with less favourable monetary outcomes (Green et al., 2012; Starns & Ratcliff, 2012; Summerfield & Koechlin, 2010; Voss et al., 2004). Our results extend these findings to moral judgement valence, suggesting that people are inclined to default to positive judgements. This may be because of the sensitivity of the bias parameter to social outcomes, such as the repercussions that come with placing moral blame improperly (Monroe & Malle, 2019; Siegel et al., 2018). Alternatively, this may also be because people expect others to behave well (Brañas-Garza, Rodríguez-Lara, & Sánchez, 2017; Yamagishi et al., 2013), and thus may also expect to select positive judgement options. Indeed, when there are prior expectations that a given response outcome is more likely, the starting point is often biased towards that response (Mulder, Wagenmakers, Ratcliff, Boekel, & Forstmann, 2012). However, we do not have insight into participants' expectations about the Decision-makers in our experiment, nor into how such expectations may change as individuals observe the full range of Decision-makers' behaviours across an experiment. Consequently, further research is needed to investigate the role of expectations in third-party moral judgements. Overall, our findings suggest that people take longer to make judgements about negative actions both because it takes them longer to process negative information, and because they favour positive judgements.

Fig. 3 suggests that there might be nonlinear patterns in the relationship between RT and offer magnitude in both experiments. Specifically, judgements of $5 offers and offers ranging $0–1 and $9–10 were generally faster than judgements of offers in the respective $3–4 and $6–7 ranges. The relatively faster responses for extremely high and extremely low offers (compared to their relatively less extreme counterparts) are well accounted for by higher absolute drift-rates for prototypical examples of morally good and bad behaviour. However, faster responses for even split offers are not entirely accounted for by drift-rate changes as absolute drift rates for $5 offers were not larger than those for $6–7 offers. Participants may have interpreted this offer as the Decision-maker's adherence to the norm of equality. Even though we explicitly excluded individuals who judged $5 offers as morally better than higher offers, the detection of equality may still create an 'anchor point', i.e., an easier case to judge (compared to the neighbouring amounts of $4 and $6), thus speeding up responses beyond what is captured by the proposed DDM model. Contrary to this idea, one might argue that an equal split creates an exceptionally difficult case to judge for the selected group of people (i.e., because equal split could be perceived as neither good nor bad). This would be expected to produce a slowing of judgements, which is not what we observed.

Our finding that people have biased caution against making negative judgements complements recent findings showing that people are more prone to adjust and change negative rather than positive beliefs about others when judging their moral character (Siegel et al., 2018). Although negative beliefs are more susceptible to change, our results suggest that people are more cautious to form these beliefs in the first place. Together, these findings suggest that people are more careful about being accurate when evaluating morally negative evidence, both in terms of changing their minds when receiving information updates (Andrejević et al., 2020; Siegel et al., 2018), and by allowing themselves time to consider all the information that is available when prompted to make a judgement. At the same time our finding that people have biased caution against making negative judgements, may not apply in situations where costs associated with incorrectly judging someone's behaviour as positive outweigh the costs associated with incorrectly judging their behaviour as negative (e.g., when responsible for selecting a person to perform a role that requires immaculate moral integrity). Indeed, previous research suggests that people require less negative information when considering suitability of a job candidate for a demanding post (Yzerbyt & Leyens, 1991), possibly because negative information is more diagnostic in these cases (Mende-Siedlecki, Baron, & Todorov, 2013; Skowronski & Carlston, 1992). The role of moral decision caution when evaluating each piece of information in such situation is an exciting avenue for future research.

Our finding that signed drift rates showed a monotonic relationship with the magnitude of Decision-makers' offers is in line with the idea that moral prototypicality of the action determines the quality of evidence for moral badness and goodness. Previous research showed that drift rate scales with perceptual discriminability of the stimuli in classical perceptual decision tasks (Voss et al., 2004). Our findings suggest that this effect generalizes to moral decisions, which is in line with the idea that moral prototypicality (i.e., how well a moral action represents adherence to or deviation from a moral norm) equates to moral discriminability and determines the rate of moral decision evidence accumulation. The monotonic relationship between the offer magnitude and the signed drift rates was, however, not linear (see Fig. 4c and d). There was an asymmetry across positive and negative offers, with mildly negative offers ($3 and $4) having a much lower absolute drift rate than mildly positive offers ($6 and $7). This may indicate that participants perceived $6 and $7 offers as much more prototypically good than they perceived offers of $3 and $4 as prototypically bad. These asymmetries are consistent with previous suggestions that people are better attuned to magnitude changes for negatively valenced moral actions and judgements of blame, as compared to positively valenced actions and judgements of praise (Anderson et al., 2020). Note however, that this theory would also predict moral judgements of positively valenced actions to be less sensitive to offer magnitude, for which we found no evidence (see Supplementary Post Hoc Analyses). Future research could study more fine-grained changes in offer size to more precisely describe the function that maps offer sizes to moral prototypicality, as well as better characterise individual difference in these mappings.

We did not find support for our hypothesis that context-expectancy would interact with the moral valence effect. Our RT results instead suggest that these two effects were additive. These results are somewhat in discord with a previous finding that some participants reduced the intensity of their negative moral judgements (but not positive moral judgements) when expecting a contextual update (Andrejević et al., 2020). There are several explanations for this discrepancy. This previous finding may be specific to moral judgements reported on a continuous scale. It may also occur only in smaller subset of people. Our strict focus on a subsample of people that condemn low offers may have excluded the people that reduce the intensity of this condemnation and show this effect. Future studies could preselect samples of people who show this effect and characterise their decision process specifically.

There are several remaining open questions that should be investigated in future studies. One outstanding question is whether the DDM can be applied to better characterise aspects of moral decision-making across a wider range of contexts. While the DDM has primarily been used to derive psychologically meaningful parameters in perceptual decision tasks (Gomez et al., 2007; Milosavljevic et al., 2010; Ratcliff, 1978; Ratcliff, Gomez, & McKoon, 2004), and has only been applied to a small range of social and more specifically moral tasks (Hutcherson et al., 2015; Pärnamets et al., 2013; Pärnamets, Richardson, & Balkenius, 2014, 2015; Son et al., 2019), our results illustrate that the DDM can be a powerful tool for dissociating parts of the decision-making process in social tasks. Our findings show that the DDM can be used to clearly partition RT variance in such tasks, and the consistency of results across two samples suggest that this partitioning is reliable. Future studies could test how well our findings generalize to other kinds of judgement tasks (e.g., traditional moral dilemmas), other moral norms (e.g., concerning harm), and other kinds of contextual information (e.g., relational status between moral actors). It could further be tested whether there is an even better model within the DDM framework to capture the process of moral judgement. Our participant exclusion criteria have restricted the sample to individuals who endorse high offers more than low offers, and who condemn lower offers. Consequently, our approach does not fully capture the diversity of moral judgement strategies that exists across individuals and cultures, and excludes those who condemned offers that were higher than an equal split (e.g., Henrich et al., 2006). This means that moral decision caution of such individuals is not accounted for in our study. We have restricted our analyses to the most plausible (and hypothesis driven) model instead of exploring the full space of all possible models, which was beyond the scope of our study. Future research, however, can extend this framework, for example by including parameters such as collapsing decision bounds (Churchland, Kiani, & Shadlen, 2008; Ditterich, 2006; Milosavljevic et al., 2010), or by allowing for inter-trial variability of some parameters (Ratcliff, 2013; Ratcliff & Rouder, 1998) to further improve the model fit; however additional theoretical work is needed to justify inclusion of such variations for the current context. Additionally, our study remains agnostic to neural mechanisms behind the moral decision process. To better understand the computation behind moral decisions, future studies should investigate the neural correlates of these computations.

To conclude, our findings identify expectancy of learning new contextual information and moral valence as impacting two distinct forms of moral decision caution. While context expectancy slows moral judgements to reduce erroneous responding in general, negatively valenced information also leads to slower judgements, presumably reducing the likelihood of making an erroneous negative judgement. Additionally, we also show that this effect of negative valence occurs in addition to another effect – that negative evidence is accumulated at a slower rate than positive evidence. These findings improve our understanding of processes underlying moral decision-making in dynamic situations and provide a foundation for future research on neural mechanisms underlying moral decisions.

Credit author statement

M.A., S.B., and D.F. contributed to conceptualization and methodology. M.A. and J.W. contributed to software, data curation, formal analysis, investigation and visualization. S.B. acquired funding. S.B., D. F., and S.L. supervised the project. M.A. administered the project and wrote the original draft. All authors reviewed and edited the manuscript.

Author note

Milan Andrejević https://orcid.org/0000-0002-7112-9384
Joshua P. White https://orcid.org/0000-0003-3129-5473
Daniel Feuerriegel https://orcid.org/0000-0002-0357-1920
Simon Laham https://orcid.org/0000-0002-9101-9553
Stefan Bode https://orcid.org/0000-0002-0258-7795

Data of all participants, materials including the instructions and the task code, as well as the analyses scripts that support the findings of this study are publicly available on an Open Science Framework (OSF) repository (DOI: 10.17605/OSF.IO/EPD63). This study was supported by an Australian Research Council grant (ARC DP160103353) to S.B. We have no competing interests. We thank Pragya Arora for her help with data collection as well as Gabriel Ong and William Turner for helpful discussions.

Correspondence concerning this article should be addressed to Milan Andrejević, Melbourne School of Psychological Sciences, Redmond Barry Building, The University of Melbourne, Parkville, Victoria 3010, Australia. Email: milan.andrejevic@unimelb.edu.au

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2022.105026.

References

Abele, A. (1985). Thinking about thinking: Causal, evaluative and finalistic cognitions about social situations. *European Journal of Social Psychology, 15*(3), 315–332. https://doi.org/10.1002/ejsp.2420150306

Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.)*, Vol. 2. Advances in experimental social psychology* (pp. 267–299). Academic Press. https://doi.org/10.1016/S0065-2601(08)60108-2.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Anderson, R. A., Crockett, M. J., & Pizarro, D. A. (2020). A theory of moral praise. *Trends in Cognitive Sciences, 24*(9), 694–703. https://doi.org/10.1016/j.tics.2020.06.008

Andrejević, M., Feuerriegel, D., Turner, W., Laham, S., & Bode, S. (2020). Moral judgements of fairness-related actions are flexibly updated to account for contextual information. *Scientific Reports, 10*(1), 17828. https://doi.org/10.1038/s41598-020-74975-0

Andrejević, M., Smillie, L. D., Feuerriegel, D., Turner, W. F., Laham, S. M., & Bode, S. (2021). How do basic personality traits map onto moral judgments of fairness-related actions? *Social Psychological and Personality Science, 194855062110382*. https://doi.org/10.1177/19485506211038295

Andrejević, M., White, J. P., Feuerriegel, D., Laham, S. M., & Bode, S. (2021). *Response time modelling reveals evidence for multiple, distinct sources of moral decision caution.* https://doi.org/10.1101/2021.01.27.427353

Baron, J., & Gürçay, B. (2017). A meta-analysis of response-time tests of the sequential two-systems model of moral judgment. *Memory & Cognition, 45*(4), 566–575. https://doi.org/10.3758/s13421-016-0686-8

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences, 107*(50), 21767–21772. https://doi.org/10.1073/pnas.0908104107

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*(4), 323–370. https://doi.org/10.1037/1089-2680.5.4.323

Berg, M. K., Kitayama, S., & Kross, E. (2021). How relationships bias moral reasoning: Neural and self-report evidence. *Journal of Experimental Social Psychology, 95*(April), 104156. https://doi.org/10.1016/j.jesp.2021.104156

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology and Evolution, 24*(3), 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Bond, K., Dunovan, K., & Verstynen, T. (2018). Value-conflict and volatility influence distinct decision-making processes. In *2018 Conference on Cognitive Computational Neuroscience* (p. 3). https://doi.org/10.32470/CCN.2018.1068-0

Brañas-Garza, P., Rodríguez-Lara, I., & Sánchez, A. (2017). Humans expect generosity. *Scientific Reports, 7*(February), 2–10. https://doi.org/10.1038/srep42446

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., … Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R Journal, 9*(2), 378–400. https://doi.org/10.32614/rj-2017-066

Churchland, A. K., Kiani, R., & Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience, 11*(6), 693–702. https://doi.org/10.1038/nn.2123

Cohen, D. J., & Ahn, M. (2016). A subjective utilitarian model of moral judgment. *Journal of Experimental Psychology: General, 145*(10), 1359–1381.

Cook, E. P., & Maunsell, J. H. R. (2002). Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nature Neuroscience, 5*(10), 985–994. https://doi.org/10.1038/nn924

De Boeck, P., & Jeon, M. (2019). An overview of models for response times and processes in cognitive tests. *Frontiers in Psychology, 10*(FEB). https://doi.org/10.3389/fpsyg.2019.00102

Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues, 31*(3), 137–149. https://doi.org/10.1111/j.1540-4560.1975.tb01000.x

Diekmann, A. (2004). The power of reciprocity: Fairness, reciprocity, and stakes in variants of the dictator game. *Journal of Conflict Resolution, 48*(4), 487–505. https://doi.org/10.1177/0022002704265948

Ditterich, J. (2006). Stochastic models of decisions about motion direction: Behavior and physiology. *Neural Networks, 19*(8), 981–1012. https://doi.org/10.1016/j.neunet.2006.05.042

Dunovan, K., & Verstynen, T. (2019). Errors in action timing and inhibition facilitate learning by tuning distinct mechanisms in the underlying decision process. *Journal of Neuroscience, 39*(12), 2251–2264. https://doi.org/10.1523/JNEUROSCI.1924-18.2019

Earp, B. D., Mcloughlin, K. L., Monrad, J. T., Clark, M. S., & Crockett, M. J. (2020). How social relationships shape moral judgment. In *, Vol. 1*. *PsyArXiv Preprints*. Issue April.

Ethridge, L. E., Brahmbhatt, S., Gao, Y., McDowell, J. E., & Clementz, B. A. (2009). Consider the context: Blocked versus interleaved presentation of antisaccade trials. *Psychophysiology, 46*(5), 1100–1107. https://doi.org/10.1111/j.1469-8986.2009.00834.x

Feather, N. T. (1999). Judgments of deservingness: Studies in the psychology of justice and achievement. *Personality and Social Psychology Review, 3*(2), 86–107. https://doi.org/10.1207/s15327957pspr0302_1

Feather, N. T., & Deverson, N. H. (2000). Reactions to a motor-vehicle accident in relation to mitigating circumstances and the gender and moral worth of the driver. *Journal of Applied Social Psychology, 30*(1), 77–95. https://doi.org/10.1111/j.1559-1816.2000.tb02306.x

Fehr, E., & Fischbacher, U. (2004). Third-party punishment and social norms. *Evolution and Human Behavior, 25*(2), 63–87. https://doi.org/10.1016/S1090-5138(04)00005-4

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics, 114*(3), 817–868. https://doi.org/10.1162/003355399556151

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology, 38*(6), 889–906. https://doi.org/10.1037/0022-3514.38.6.889

Forstmann, B. U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wagenmakers, E. J., … Turner, R. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences of the United States of America, 107*(36), 15916–15920. https://doi.org/10.1073/pnas.1004932107

Gallotti, R., & Grujić, J. (2019). A quantitative description of the transition between intuitive altruism and rational deliberation in iterated Prisoner's Dilemma experiments. *Scientific Reports, 9*(1), 1–11. https://doi.org/10.1038/s41598-019-52359-3

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience, 30*(1), 535–574. https://doi.org/10.1146/annurev.neuro.29.051605.113038

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General, 136*(3), 389–413. https://doi.org/10.1037/0096-3445.136.3.389

Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology, 96*(5), 1029–1046. https://doi.org/10.1037/a0015141

Green, N., Biele, G. P., & Heekeren, H. R. (2012). Changes in neural connectivity underlie decision threshold modulation for reward maximization. *Journal of Neuroscience, 32*(43), 14942–14950. https://doi.org/10.1523/JNEUROSCI.0573-12.2012

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization, 3*(4), 367–388. https://doi.org/10.1016/0167-2681(82)90011-7

Haidt, J., & Baron, J. (1996). Social roles and the moral judgement of acts and omissions. *European Journal of Social Psychology, 26*(2), 201–218. https://doi.org/10.1002/(SICI)1099-0992(199603)26:2<201::AID-EJSP745>3.0.CO;2-J

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., … Ziker, J. (2006). Costly punishment across human societies. *Science, 312*(5781), 1767–1770. https://doi.org/10.1126/science.1127333

Ho, A. K., Sidanius, J., Kteily, N., Sheehy-Skeffington, J., Pratto, F., Henkel, K. E., … Stewart, A. L. (2015). The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO_7 scale. *Journal of Personality and Social Psychology, 109*(6), 1003–1028. https://doi.org/10.1037/pspi0000033

Hutcherson, C. A., Bushong, B., & Rangel, A. (2015). A neurocomputational model of altruistic choice and its implications. *Neuron, 87*(2), 451–463. https://doi.org/10.1016/j.neuron.2015.06.031

Hysom, S. J., & Fişek, M. H. (2011). Situational determinants of reward allocation: The equity-equality equilibrium model. *Social Science Research, 40*(4), 1263–1285. https://doi.org/10.1016/j.ssresearch.2011.02.002

Kiani, R., & Shadlen, M. N. (2009). Representation of confidence associated with a decision by neurons in the parietal cortex. *Science, 324*(5928), 759–764. https://doi.org/10.1126/science.1169405

Kleiner, M., Brainard, D. H., Pelli, D. G., Broussard, C., Wolf, T., & Niehorster, D. (2007). What's new in Psychtoolbox-3? *Perception*. https://doi.org/10.1068/v070821

Krajbich, I., Hare, T., Bartling, B., Morishima, Y., & Fehr, E. (2015). A common mechanism underlying food choice and social decisions. *PLoS Computational Biology, 11*(10), 1–24. https://doi.org/10.1371/journal.pcbi.1004371

Lee, K., & Ashton, M. C. (2018). Psychometric properties of the HEXACO-100. *Assessment, 25*(5), 543–556. https://doi.org/10.1177/1073191116659134

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6*(August), 1–16. https://doi.org/10.3389/fpsyg.2015.01171

Maris, E. (1993). Additive and multiplicative models for gamma distributed random variables, and their application as psychometric models for response times. *Psychometrika, 58*(3), 445–469. https://doi.org/10.1007/BF02294651

Meade, B., Lafayette, L., Sauter, G., & Tosello, D. (2017). *Spartan HPC-Cloud hybrid: Delivering performance and flexibility.*. https://doi.org/10.4225/49/58ead90dceaaa

Mende-Siedlecki, P., Baron, S. G., & Todorov, A. (2013). Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *Journal of Neuroscience, 33*(50), 19406–19415. https://doi.org/10.1523/JNEUROSCI.2334-13.2013

Meristo, M., & Surian, L. (2013). Do infants detect indirect reciprocity? *Cognition, 129*(1), 102–113. https://doi.org/10.1016/j.cognition.2013.06.006

Messick, D. M., & Schell, T. (1992). Evidence for an equality heuristic in social decision making. *Acta Psychologica, 80*(1–3), 311–323. https://doi.org/10.1016/0001-6918(92)90053-G

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision making, 5*(6), 437–449. https://doi.org/10.2139/ssrn.1901533

Miron, A. M., Warner, R. H., & Branscombe, N. R. (2011). Accounting for group differences in appraisals of social inequality: Differential injustice standards. *British Journal of Social Psychology, 50*(2), 342–353. https://doi.org/10.1111/j.2044-8309.2010.02009.x

Monroe, A. E., & Malle, B. F. (2019). People systematically update moral judgments of blame. *Journal of Personality and Social Psychology, 116*(2), 215–236. https://doi.org/10.1037/pspa0000137

Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., & Forstmann, B. U. (2012). Bias in the brain: A diffusion model analysis of prior probability and potential payoff. *Journal of Neuroscience, 32*(7), 2335–2343. https://doi.org/10.1523/JNEUROSCI.4156-11.2012

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature, 393*(6685), 573–577. https://doi.org/10.1038/31225

Olson, J. G., McFerran, B., Morales, A. C., & Dahl, D. W. (2016). Wealth and welfare: Divergent moral reactions to ethical consumer choices. *Journal of Consumer Research, 42*(6), 879–896. https://doi.org/10.1093/jcr/ucv096

Palmer, E. M., Horowitz, T. S., Torralba, A., & Wolfe, J. M. (2011). What are the shapes of response time distributions in visual search? *Journal of Experimental Psychology: Human Perception and Performance, 37*(1), 58–71. https://doi.org/10.1037/a0020747

Pärnamets, P., Johansson, P., Balkenius, C., Hall, L., Spivey, M. J., & Richardson, D. C. (2013). Changing minds by tracking eyes: Dynamical systems, gaze and moral decisions. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 1115–1120.

Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J., & Richardson, D. C. (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy of Sciences of the United States of America, 112*(13), 4170–4175. https://doi.org/10.1073/pnas.1415250112

Pärnamets, P., Richardson, D., & Balkenius, C. (2014). Modelling moral choice as a diffusion process dependent on visual fixations. In *Proceedings of the Annual Meeting of the Cognitive Science Society*.

Piazza, J., & Sousa, P. (2013). Religiosity, political orientation, and consequentialist moral thinking. *Social Psychological and Personality Science, 5*(3), 334–342. https://doi.org/10.1177/1948550613492826

Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *The Journal of Neuroscience, 27*(37), 9984–9988. https://doi.org/10.1523/JNEUROSCI.2131-07.2007

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R. (2013). Parameter variability and distributional assumptions in the diffusion model. *Psychological Review, 120*(1), 281–292. https://doi.org/10.1037/a0030775

Ratcliff, R., Gomez, P., & McKoon, G. (2004). A diffusion model account of the lexical decision task. *Psychological Review, 111*(1), 159–182. https://doi.org/10.1037/0033-295X.111.1.159

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science, 9*(5), 347–356. https://doi.org/10.1111/1467-9280.00067

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences, 20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Ratcliff, R., Thapar, A., & McKoon, G. (2004). A diffusion model analysis of the effects of aging on recognition memory. *Journal of Memory and Language, 50*(4), 408–424. https://doi.org/10.1016/j.jml.2003.11.002

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin and Review, 9*(3), 438–481. https://doi.org/10.3758/BF03196302

Sawaoka, T., Newheiser, A.-K., & Dovidio, J. F. (2014). Group-based biases in moral judgment: The role of shifting moral standards. *Social Cognition, 32*(4), 360–380. https://doi.org/10.1521/soco.2014.32.4.360

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Shapiro, E. G. (1975). Effect of expectations of future interaction on reward allocations in dyads: Equity or equality. *Journal of Personality and Social Psychology, 31*(5), 873–880. https://doi.org/10.1037/h0076703

Siegel, J. Z., Mathys, C., Rutledge, R. B., & Crockett, M. J. (2018). Beliefs about bad people are volatile. *Nature Human Behaviour, 2*(10), 750–756. https://doi.org/10.1038/s41562-018-0425-1

Simpson, A., Laham, S. M., & Fiske, A. P. (2016). Wrongness in different relationships : Relational context effects on moral judgment. *The Journal of Social Psychology, 156* (6), 594–609. https://doi.org/10.1080/00224545.2016.1140118
Skowronski, J. J., & Carlston, D. E. (1992). Caught in the act: When impressions based on highly diagnostic behaviours are resistant to contradiction. *European Journal of Social Psychology, 22*(5), 435–452. https://doi.org/10.1002/ejsp.2420220503
Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences, 27*(3), 161–168. https://doi.org/10.1016/j.tins.2004.01.006
Son, J. Y., Bhandari, A., & Feldman Hall, O. (2019). Crowdsourcing punishment: Individuals reference group preferences to inform their own punitive decisions. *Scientific Reports, 9*(1), 1–15. https://doi.org/10.1038/s41598-019-48050-2
Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society. Series B: Statistical Methodology, 64*(4), 583–616. https://doi.org/10.1111/1467-9868.00353
Starns, J. J., & Ratcliff, R. (2012). Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychonomic Bulletin and Review, 19*(1), 139–145. https://doi.org/10.3758/s13423-011-0189-3
Stoet, G., & Snyder, L. H. (2007). Extensive practice does not eliminate human switch costs. *Cognitive, Affective, & Behavioral Neuroscience, 7*(3), 192–197. https://doi.org/10.3758/CABN.7.3.192
Summerfield, C., & Koechlin, E. (2010). Economic value biases uncertain perceptual choices in the parietal and prefrontal cortices. *Frontiers in Human Neuroscience, 4* (November), 1–12. https://doi.org/10.3389/fnhum.2010.00208
Turiel, E. (2001). *The culture of morality*. Cambridge University Press.. https://doi.org/10.1017/CBO9780511613500
Van Zandt, T. (2002). Analysis of response time distributions. In *Stevens' handbook of experimental psychology* https://doi.org/10.1002/0471214426.pas0412
Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory and Cognition, 32*(7), 1206–1220. https://doi.org/10.3758/BF03196893
Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods, 17*(2), 228–243. https://doi.org/10.1037/a0027127
Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin and Review, 11*(1), 192–196. https://doi.org/10.3758/BF03206482
Weidman, A. C., Sowden, W. J., Berg, M. K., & Kross, E. (2020). Punish or protect? How close relationships shape responses to moral violations. *Personality and Social Psychology Bulletin, 46*(5), 693–708. https://doi.org/10.1177/0146167219873485
Wentura, D., Rothermund, K., & Bak, P. (2000). Automatic vigilance: The attention-grabbing power of approach- and avoidance-related social information. *Journal of Personality and Social Psychology, 78*(6), 1024–1037. https://doi.org/10.1037/0022-3514.78.6.1024
Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical bayesian estimation of the drift-diffusion model in Python. *Frontiers in Neuroinformatics, 7*(JULY 2013), 1–10. https://doi.org/10.3389/fninf.2013.00014
Yamagishi, T., Mifune, N., Li, Y., Shinada, M., Hashimoto, H., Horita, Y., Miura, A., Inukai, K., Tanida, S., Kiyonari, T., Takagishi, H., & Simunovic, D. (2013). Is behavioral pro-sociality game-specific? Pro-social preference and expectations of pro-sociality. *Organizational Behavior and Human Decision Processes, 120*(2), 260–271. https://doi.org/10.1016/j.obhdp.2012.06.002
Yu, H., Siegel, J. Z., Clithero, J. A., & Crockett, M. J. (2021). How peer influence shapes value computation in moral decision-making. *Cognition, 211*(February), Article 104641. https://doi.org/10.1016/j.cognition.2021.104641
Yzerbyt, V. Y., & Leyens, J. P. (1991). Requesting information to form an impression: The influence of valence and confirmatory status. *Journal of Experimental Social Psychology, 27*(4), 337–356. https://doi.org/10.1016/0022-1031(91)90030-A