

To appear in British Journal of Psychology (accepted 27th September 2021)

They Should Have Known Better: The Roles of Negligence and Outcome in Moral Judgments of
Accidental Actions

Gavin Nobes

University of East Anglia

Justin W. Martin

Boston College

Author note

Gavin Nobes, School of Psychology, University of East Anglia, Norwich, UK; Justin Martin, Department of Psychology and Neuroscience, Boston College, MA.

Correspondence concerning this article should be addressed to Gavin Nobes, School of Psychology, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom. Email: g.nobes@uea.ac.uk

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Abstract

Two experiments were conducted to investigate the relative influence of agents' negligence and their actions' unintended outcomes on moral judgments. In Study 1, 343 participants were asked in an online questionnaire about a driver whose level of negligence, and the severity of the outcome, were varied systematically. Each judged how much punishment and blame the driver deserved, and rated her negligence, causal responsibility, and intentionality. In Study 2, 341 participants completed the same questionnaire, and also judged the driver's wrongness and the outcome's severity. In both studies, judgments were strongly influenced by negligence; blame was also affected by causal responsibility, and wrongness by intention, but the relatively slight outcome effect on blame and wrongness was largely mediated by negligence. In contrast, both negligence and outcome had substantial effects on punishment judgments: most participants assigned high levels of punishment when, and only when, the outcome was negative *and* the agent was negligent. These findings indicate that the processes underlying punishment judgments differ from those involved in blame and wrongness judgments, and when no negligence information is provided in the description of accidents (as in many previous studies), participants often attribute negligence to agents, and judge them accordingly. The findings also suggest that the effect of outcome on moral judgments has often been overestimated by researchers, and that of negligence underestimated.

Keywords: Moral judgments; Negligence; Outcome; Moral luck; Punishment; Blame

They Should Have Known Better: The Roles of Negligence and Outcome in Moral Judgments of
Accidental Actions

One fall afternoon, Cynthia is driving home from getting lunch. Her two brothers are in the car with her. As they get close to their home, they see a very large leaf pile on the side of the road, the biggest they have seen all season. Cynthia's younger brother asks her to drive through the leaf pile, and she agrees. Cynthia drives the car through the leaf pile, and she feels two bumps as the car passes over some objects. Unknown to Cynthia, two children were hiding in the leaf pile. The children were hit by the car and killed instantly.

When Martin and Cushman (2016) presented this vignette to online participants, 94% said that the driver should be punished, and the average sentence was 1-3 years in prison. In contrast, when the same story finished with "Unknown to Cynthia, there were some large sticks in the leaf pile. The car was not damaged by the sticks", 85% assigned no punishment. Clearly, there was a strong outcome effect: all other aspects of the story were identical, and so people's moral judgments must have been influenced by the consequences of Cynthia's actions.

This stark finding and others (e.g., Gino et al., 2009; Mazzocco et al., 2004; Robbennolt, 2000; Walster, 1966) provide examples of the intriguing phenomenon of moral luck, which has attracted considerable debate among philosophers for several decades (Hartman, 2017; Nagel, 1979; Nelkin, 2021; Otsuka, 2009; Williams, 1981), and, more recently, experimental investigation by moral psychologists (e.g., Kneer & Machery, 2019; Lench et al., 2015; Martin & Cushman, 2016; Young et al., 2010). In Martin and Cushman's vignettes, Cynthia was lucky

when there were only sticks in the leaf pile and so received little or no punishment, and unlucky when she ran over the children, and was assigned much harsher punishment.

Moral luck is puzzling because fair and rational judgments should be based on morally-relevant factors such as intentions and beliefs: According to the Control Principle (Kant, 1784 / 1998), “Two people ought not to be morally assessed differently if the only other differences between them are due to factors beyond their control.” (Nelkin, 2021). However, apparently morally-irrelevant factors such as outcomes and luck seem to play important roles in our moral judgments, and Martin and Cushman’s (2016) findings illustrate how, in practice, we often assess people according to factors – such as children hiding in a leaf pile – that are entirely beyond their control.

This strong effect of outcome on moral judgments is also puzzling because a well-established finding in developmental psychology is that there is an “outcome-to-intent shift” such that, with age, children increasingly base their moral judgments on intentions rather than outcomes (e.g., Costanzo et al., 1973; Cushman et al., 2013; Farnill, 1974; Piaget, 1932; Zelazo et al., 1996). Moreover, there is now a strong body of evidence that indicates that younger children, and even infants, show a surprisingly sophisticated awareness of intention information, and that they can and often do prioritize this over outcome information in their moral judgments (e.g., Dunfield & Kuhlmeier, 2010; Hamlin, 2013; Margoni & Surian, 2016, 2020; Nobes et al., 2009; Van de Vondervoort & Hamlin, 2018). It follows that adults’ judgments of Cynthia should be based on her benign intentions – she did not mean to cause harm – and that the outcome should be irrelevant, so it is intriguing that this is not the case.

Martin and Cushman’s explanation of their findings derives from the two-process model of moral judgment proposed by Cushman and colleagues (Cushman, 2008; Cushman et al., 2013). According to this model, adults’ wrongness judgments (i.e., whether an agent or action is morally good or bad) are influenced by a process that evaluates the agent’s intentions and

beliefs. In contrast, punishment and blame judgments are influenced both by this mental state process and by a second process that assesses the agent's causal role. In cases of accidental harm like Cynthia's, when the agent had no intention to harm, participants' punishment and blame judgments are influenced by a competitive interaction between the causal process, which inculpates based upon the degree of harm caused, and the mental state process, which exculpates based upon the agent's benign intent. The two-process model therefore leads to the predictions that, regardless of the outcome, Cynthia will not be considered wrong¹ because her intention was benign; neither will she be assigned blame or punishment when no harm is done (the sticks are broken); but that she will be assigned blame and punishment when the outcome is negative (the children are killed).

But there are other factors that influence our moral judgments besides intention and outcome. One such is negligence (e.g., Enzle & Hawkins, 1992; Laurent et al., 2016; Margoni et al., 2019; Margoni & Surian, 2021; Nuñez et al., 2014; Rizzo et al., 2019; Shultz & Wright, 1985). In the Cynthia vignettes, participants might have been influenced not only by her benign intentions, or by the positive or negative outcome, but also by their assumptions about how she was driving – perhaps dangerously fast, or off the road – or about what she *should* have been aware of, such as that children might be hiding in the leaf pile. Like intention, negligence is morally relevant, and this is reflected in the law: drivers who cause harm through driving without due care and attention are blameworthy and punishable, even when there is no suggestion that they intended the outcome.

Recently, researchers have explored whether assumed negligence might explain young children's apparent tendency to judge agents according to the outcomes of their actions, even if

¹ While blame and punishment are ascribed to agents, moral wrongness is typically considered a feature of their actions. However, action-focused wrongness questions can lead people to assess the *outcomes* of actions, and therefore to give misleadingly outcome-based judgments. In practice, many moral psychologists have asked about the wrongness of *agents* for acting as they did. See Nobes et al. (2016) for a discussion and report of the effect of rephrasing wrongness questions in this way.

they were well-intentioned (Killen et al., 2011; Nobes et al., 2017; Nobes et al., 2009). After all, even very young children are frequently told to be careful, and they are admonished when they accidentally cause harm through carelessness. For example, when considering a well-intentioned girl who accidentally dropped her pet, children might assume that she was punishable not because she hurt the pet (the outcome), but because she didn't hold it carefully enough (negligence). Indeed, when they were told explicitly that the accident occurred despite the girl being careful, children were found to be more lenient, and to make more 'mature', intention-based judgments (Nobes et al., 2017). Moreover, Mulvey et al. (2020) report that young children are sensitive to the negligence of victims, as well as that of agents.

Building on this recent developmental research, we report here two experiments in which we investigated the possibility that adults' judgments of unintended (accidental) actions might also be strongly influenced by negligence rather than by outcome. If so, then Martin and Cushman's findings, and those of other researchers who have reported outcome effects in similar circumstances (e.g., Gino et al., 2010; Patil et al., 2017; Robbennolt, 2000; Walster, 1966), might be accounted for in terms of participants considering unlucky agents – for whom outcomes are negative – more negligent than lucky agents, for whom they are positive. Judgments would therefore co-vary with outcome and so seem outcome-based, but would actually be based on perceived negligence. If this negligence-based explanation were correct, then controlling for negligence should reduce or remove the outcome effect on moral judgments.

There is some evidence for this being the case. Mazzocco et al. (2004) found outcome effects on blame attributions, but that these were largely mediated by perceived negligence: agents were considered more negligent when the outcome was negative. Similarly, Kneer and Machery (2019) reported that the outcome effects on blame and wrongness judgments were accounted for by the agent being considered more negligent when the outcome was negative. However, this mediation by negligence only partially explained the outcome effect on

punishment judgments: participants assigned punishment both according to the agent's negligence and to the severity of the outcome per se.

The current studies. We examined these possible explanations of moral luck by replicating Martin and Cushman's (2016) experiment, and manipulating the agent's level of negligence. Since Cynthia's intentions remained constant and benign, the two-process model predicts that outcome would be the dominant influence on blame and punishment judgments, and that adding and varying negligence information should make little or no difference.

On the other hand, if moral judgments of accidental harms are accounted for primarily by perceived negligence, two key predictions follow: first, the addition of negligence information should strongly and directly influence participants' moral judgments so that negligent agents are condemned more, and non-negligent agents less, independent of the outcome. And second, the outcome effect – and hence moral luck – should be explained by participants perceiving the driver to be more negligent when the children are killed than when no harm is done.

We varied negligence in four ways: first, no explicit negligence information was given, as in the original study. The second, third and fourth levels of this negligence IV followed the wording used in vignettes by Shen et al. (2011) to distinguish their blameless (non-negligent), negligent and reckless actions (see Table 1 for the phrasing of each of these levels). Based on the US Model Penal Code (1962), an agent is described as negligent when they “should be aware of a substantial and unjustifiable risk”, whereas a reckless agent is one who “consciously disregards a substantial and unjustifiable risk”, (p. 1316).

In addition to punishment judgments, we asked participants to assess the driver's blameworthiness and negligence. The latter allowed us to assess both the validity of our manipulation of the negligence IV and the extent to which perceived negligence was associated with punishment and blame judgments. We also asked participants to rate the driver's causal responsibility for the accident, and the extent to which she intended the outcome.

The two-process model (Cushman et al., 2008; 2013) and the negligence-based account outlined above led to the contrasting predictions about punishment and blame judgments shown in Table 2 (preregistration <https://aspredicted.org/blind.php?x=uk8em3>, reference #22404).

Study 1. Method

Design. We used a 4 (Negligence IV [original (no negligence information), non-negligent, negligent, reckless]) x 2 (Outcome IV [no harm, children killed]) between-subjects design. The dependent variables were the two moral judgments – deserved punishment and blame – and the three covariates were perceived negligence, causal responsibility, and intention. All five of these variables were measured on 0-10 Likert scales.

Sample. Participants were recruited from a psychology student panel and through social media. The students received course credits for participation. Only respondents aged 18 and over took part. Of the 358 who started the questionnaire, 13 (3.6%) were excluded because they asked for their data not to be used, as were two (0.6%) others because their responses to open questions indicated poor engagement and understanding. The sample size was therefore 343. A priori power analyses using G*Power 3.1 (Faul et al., 2007) indicated that, with power set at .95 and $\alpha = .05$, for the ANOVAs a sample of $N = 279$ (i.e., 35 per group) would be sufficient to detect a medium effect size of $f = 0.25$ ($\eta_p^2 = .06$), and for the multiple regressions a sample of $N = 138$ would be sufficient to detect a medium effect size of $f^2 = 0.15$.

The mean age of the participants was 28.7 (SD = 12.1) years; 240 (70.0%) were women; 83 (24.2%) were psychology undergraduates; and 104 (30.3%) were university graduates. 280 (81.6%) described themselves as white, and English was the first language of 298 (86.9%).

Materials and procedure. The eight versions of the vignette are shown in Table 1, and the questions and 11-point rating scales in Table 3. The invitations included links to an online Qualtrics questionnaire. Following information about the study and a consent form, participants

were asked about their age, gender, education and language. The vignette and questions were then presented.

Ethics. Ethical approval was obtained from the University of X's School of Psychology Research Ethics Committee, Reference 2019-0024-001505.

Data analysis. For each of the judgment types – punishment and blame – we first tested Hypotheses 1 and 2, which concern the effects on judgments of outcome and negligence, respectively. We focused first on the impact of the outcome and negligence IVs and conducted a 2 (Outcome IV [no harm, children killed]) x 4 (Negligence IV [original, non-negligent, negligent, reckless]) between-subjects ANOVA with the judgment as the DV. Post hoc Bonferroni tests assessed differences between negligence IV levels. We then investigated the relative influence of the outcome IV and perceived negligence (as opposed to the negligence IV) by running a multiple regression for each judgment type in which outcome, negligence, and the outcome-negligence interaction were predictors of the judgment. The perceived negligence scores were mean-centred. Bias corrected and accelerated 95% confidence intervals were based on 1000 bootstrap samples. The maximum correlation (between perceived negligence and causation) between predictors was .69, and the maximum variance inflation factor (VIF) was 2.08, indicating that multicollinearity was unlikely to have impacted the models. With seven exceptions that were excluded (six from the punishment model, one from blame), the standardized residuals ranged between -3.29 and 3.29, indicating that the models were not biased by any individual cases.

Hypotheses 3 and 4 concern the possible mediation of outcome effects by negligence, and of negligence by intention and causal responsibility. These were tested by conducting mediation analyses by means of Hayes' (2017) process method, which uses ordinary least squares path analysis, and calculates 95% confidence intervals based on 5,000 bootstrap samples.

Hypothesis 5 – that causal responsibility and intention would also predict judgments – was addressed by including these factors in the multiple regression and mediation analyses.

Hypothesis 6 was tested by comparing the patterns of findings for punishment and blame to assess whether they were generally similar; if so, this would suggest that these judgments are based on similar processes.

Having investigated the influences on each judgment type, the impact of the outcome and negligence IVs on perceived negligence was investigated by running the equivalent 2 (Outcome IV) x 4 (Negligence IV) ANOVA with perceived negligence as the DV.

Results

Punishment judgments. When no negligence information was given (as in the original study), and the outcome was positive (sticks were broken), 36 (85.7%) of 42 participants said the driver should receive no punishment, and none thought she should receive more than a small fine or probation (Figure 1a). In contrast, when the outcome was negative (the children were killed), three (7.5%) of 40 participants said she should not be punished, and 30 (75.0%) assigned at least a year in jail.

This strong outcome effect on punishments was evident at all levels of the negligence IV and from the high correlation, $r = .72, p < .001$ (Table 4). However, punishment judgments were most severe when both negligence and outcome were negative: when the driver was negligent or reckless but no harm was caused, four out of 84 (4.7%) participants assigned at least a few weeks in jail, as did 14 out of 41 (34.1%) when she was not negligent but the children were killed. But of the 89 participants for whom the driver was negligent or reckless and outcome was negative, 71 (79.8%) thought she should be jailed for at least a few weeks.

An ANOVA corroborated the strong outcome effect on punishment judgments, $F(1, 335) = 430.68, p < .001, \eta_p^2 = .56$, and a smaller effect of the negligence IV, $F(3, 335) = 19.30, p < .001, \eta_p^2 = .15$. When the driver was not negligent, punishment was lower than at all other

levels, $ps < .001$, and when no negligence information was given (the original condition), she was assigned less punishment than when she was reckless, $p = .01$. However, there was no significant difference between punishment assigned in the original and negligent conditions, $p = 1.00$, nor in the negligent and reckless conditions, $p = .26$.

There was also an interaction between the negligence IV and outcome, $F(3, 335) = 11.03$, $p < .001$, $\eta_p^2 = .09$: when the outcome was negative, the influence on punishment judgments of the negligence IV, $F(3, 166) = 15.65$, $p < .001$, $\eta^2 = .22$, was greater than when the outcome was positive, $F(3, 169) = 5.25$, $p = .002$, $\eta^2 = .09$. At both levels of outcome less punishment was assigned when the driver was not negligent than when she was negligent or reckless, $ps < .03$; and there was no difference between the negligent and reckless groups' assignment of punishment ($ps > .46$).

Multiple regression also indicated a substantial effect of outcome on punishment judgments, a strong effect of perceived negligence, and an interaction between these predictors. There was no discernible effect of causation or intention on punishment judgments (Table 5a).

Mediation analysis also showed that the total effect of outcome on punishment was high, $c = 5.35$, $p < .001$, 95% CI [4.79, 5.91]; that is, when the children were killed, participants assigned on average more than five points more punishment than when there were only sticks in the leaf pile (Figure 2). The direct effect of outcome was substantial, $c' = 4.54$, $p < .001$, 95% CI [4.01, 5.07]. There was also a modest indirect effect of outcome through perceived negligence, $a_1*b_1 = 0.64$, 95% CI [0.30, 1.04]. The total effect of perceived negligence on punishment judgments, $c = 0.55$, $p < .001$, 95% CI [0.46, 0.64], was wholly accounted for by its direct effect, $c' = 0.60$, $p < .001$, 95% CI [0.48, 0.72].

Blame judgments. Correlations (Table 4) and means (Figure 1b) indicate that the effect of outcome on blame was considerably less than on punishment, and that the influence of negligence was much greater. This was corroborated by an ANOVA which indicated that blame

judgments were influenced primarily by the level of the negligence IV, $F(3, 333) = 49.16$, $p < .001$, $\eta_p^2 = .31$, and only slightly by outcome, $F(1, 333) = 6.00$, $p = .015$, $\eta_p^2 = .02$. The interaction between outcome and negligence IVs did not approach significance, $F(3, 333) = 0.84$, $p = .47$, $\eta_p^2 = .01$. Post hoc tests indicated that, while the driver was considered significantly less blameworthy in the non-negligent condition than in the other conditions, $ps < .001$, and in the original condition than when she was reckless, $p = .009$, there was no difference in blame judgments between the negligent and original conditions, $p = .33$, nor between when she was negligent and reckless, $p = 1.00$.

Multiple regression gave no indication of outcome independently influencing blame judgments (Table 5b). Instead, blame was assigned according to how negligent, and, to a lesser extent, how causally responsible the driver was considered to be. Intention ratings did not independently predict blame scores.

Mediation analysis also indicated that the total effect of outcome on blame judgments was slight, $c = 0.94$, $p = .016$, 95% CI [0.18, 1.71], and there was no evidence of it having a direct effect, $c' = 0.13$, $p = .64$, 95% CI [-0.39, 0.64]. The indirect effect of outcome through negligence, $a_1*b_1 = 0.78$, 95% CI [0.36, 1.23], and, less so, through causation, $a_3*b_3 = .28$, 95% CI [0.03, 0.57] accounted for the total effect; that is, when negligence and causation were held constant, the outcome effect on blame judgments was negligible or nil.

The total effect of negligence on blame, $c = 0.77$, $p < .001$, 95% CI [0.70, 0.84], was primarily accounted for by its direct effect, $c = 0.53$, $p < .001$, 95% CI [0.44, 0.62], but there was also an indirect effect through causation, $a_3*b_3 = 0.22$, 95% CI [0.14, 0.32].

The total effect of causation on blame, $c = 0.74$, $p < .001$, 95% CI [0.67, 0.82] was partly explained by its direct effect, $c' = .34$, $p < .001$, 95% CI [0.25, 0.43], but also by its indirect effect through negligence, $a_1*b_1 = 0.39$, 95% CI [0.30, 0.47].

Perceived negligence. Participants' ratings of the driver's negligence were strongly correlated with the manipulated negligence IV, which suggests that this IV was effective, and that the perceived negligence variable was valid. Perceived negligence was also strongly associated with blame judgments, causal responsibility and, less so, punishment judgments. There was a correlation too between perceived negligence and outcome; that is, the driver was considered more negligent when she killed the children.

ANOVA indicated a strong main effect of the negligence IV, $F(3, 333) = 71.49, p < .001, \eta_p^2 = .39$ (Figure 1c). Perceived negligence was significantly lower in the non-negligent condition than in all others, $ps < .001$, and in the original condition than the reckless, $p = .007$. However, when the driver was negligent, perceived negligence did not differ significantly from the original condition, $p = .30$, nor from when she was reckless, $p = 1.00$.

There was also a main effect of outcome on perceived negligence, $F(1, 333) = 19.04, p < .001, \eta_p^2 = .05$: the driver was considered more negligent when the children were killed. The interaction between the negligence IV and outcome did not approach significance, $F(3, 333) = 0.95, p = .42, \eta_p^2 = .01$

Study 1 Discussion

Consistent with previous studies (e.g., Cushman, 2008; Gino et al., 2009; Kneer & Machery, 2019; Lench et al., 2015; Martin & Cushman, 2016; Mazzocco et al., 2004), the findings of this experiment indicate a strong outcome effect on punishment judgments. In particular, irrespective of the level of negligence, almost no participants assigned a jail sentence when the outcome was positive. However, a previously unreported finding was that the influence of outcome was moderated by negligence such that, when the outcome was negative, participants were considerably more likely to assign high levels of punishment when the driver was negligent or reckless than when she was not negligent. These results support the two-process model's first prediction that punishment judgments would be influenced substantially by

outcome. They also support the negligence-based account's second prediction that negligence would be a strong influence on punishment. Neither account predicted the strong interaction between outcome and negligence.

Blame judgments were strongly influenced by negligence, but, in contrast to punishment judgments and the findings of most previous research, (e.g., Cushman, 2008; Gino et al., 2009; Lench et al., 2015; Mazzocco et al., 2004; for a review, see Robbennolt, 2000) there was relatively little outcome effect. This is consistent with the negligence-based account's first and second predictions, and not with the two-process model's.

Also consistent with the negligence-based account's second prediction, participants considered the driver in the original condition (i.e., the absence of negligence information) to be approximately as negligent, punishable and blameworthy as in the negligent condition.

The negligence-based account's third prediction was that any outcome effect on judgments would be mediated by negligence, while the two-process model's fourth prediction was that any negligence effect would be mediated by intention and / or causal responsibility. Regarding punishment judgments, neither was supported, indicating that outcome and negligence were largely independent, though interacting, predictors. In contrast, the mediation analyses of blame judgments supported the negligence-based account because the relatively slight outcome effect was largely accounted for by its indirect effect through negligence; it occurred primarily because participants considered the driver to be more negligent when the outcome was negative than when it was positive. There was less support for the two-process model: the total effect of negligence was only partially mediated by causal responsibility.

The remaining predictions were shared by both the two-process model and the negligence-based account. The fifth was that causal responsibility and intention would influence judgments. There was no evidence of either factor influencing punishment judgments, but causal responsibility had a moderately strong influence on blame, which was only partially mediated by

negligence. The sixth prediction was that the patterns of punishment and blame judgments would be similar, and that these judgments would be similarly influenced by outcome and negligence, but in both ways they differed markedly.

In summary, as regards punishment, there was partial support for both the two-process model and the negligence-based account because both outcome and negligence were influential: the strong outcome effect was moderated by negligence such that, when the children were killed, most participants assigned a high level of punishment if, and only if, they considered the driver negligent. Regarding blame, the negligence-based account was strongly supported because these judgments were influenced substantially by negligence and, to a lesser extent, by causal responsibility. The two-process model's prediction that outcome would be the principal influence on blame was not supported.

Study 2

Study 2 was conducted to replicate Study 1, and to address the following issues. First, a possible criticism of Study 1 is that the only measure of outcome was the binary IV (sticks broken or children killed). Since it is presumably the participants' *perceptions* of outcomes that directly influence their judgments, in Study 2 we asked participants to rate the severity of the outcome using a 0-10 Likert-type scale.

Another possible problem is that, to replicate Martin and Cushman's (2016) study, the punishment question was always asked first (they did not ask about blame). The unexpected finding that punishment judgments were strongly related to outcome, but blame judgments were not, could have resulted from an order effect. In Study 2 the order of the judgment questions was therefore randomized.

In Study 1 there were no substantive differences in negligence and judgment ratings between when the driver was negligent and reckless. In Study 2 we therefore removed the reckless level of the negligence IV.

According to the two-process model, wrongness judgments are based solely on mental states, in particular intentions. Since the driver did not intend the negative outcome, the two-process model would lead to the prediction that the driver would not be considered wrong, even when the children were killed. To test this prediction, we added wrongness as a third type of judgment.

The evidence from Study 1 indicated that participants' perceptions of the driver's intention had little or no bearing on their moral judgments. However, it is possible that some, perhaps many, participants misinterpreted the question – “To what extent did Cynthia intend things to turn out the way they did?” – because it seems to focus on the outcome, rather than the driver. After all, it makes sense to say that the driver had absolutely no intention to kill the children, regardless of how or why she drove through the leaf-pile. A better question would focus on the agent herself, rather than on the outcome of her actions, and this might lead more participants to report that, for example, the driver's negligent actions were more intentional, or at least *less unintentional*, than her non-negligent actions. In this second study we therefore changed the wording of the intention question so that it was more clearly focused on the agent: “To what extent did Cynthia intend to harm anyone?”

The predictions were as in Study 1, with the addition of Hypothesis 7: the two-process model leads to the prediction that wrongness judgments are based mainly on perceptions of the driver's intention, whereas according to the negligence-based account, perceived negligence is the primary influence.

Method

Recruitment and sample. Members of two university research participant panels were invited to take part, of whom 178 started the questionnaire. Four (2.2%) chose not to have their data included, and six (3.4%) more were excluded because they gave responses that indicated misunderstanding or disengagement. In addition, 191 respondents living in the US were

recruited through Amazon's Mechanical Turk, of whom three (1.6%) opted out, and 15 (7.9%) misunderstood or failed to engage. The characteristics of the 341 participants are shown in Table 6.

A priori power analyses using G*Power 3.1 (Faul et al., 2007) indicated that, with power set at .95 and $\alpha = .05$, for the ANOVAs a sample of $N = 251$ (i.e., 42 per group) would be sufficient to detect a medium effect size of $f = 0.25$ ($\eta_p^2 = .06$), and for the multiple regressions a sample of $N = 138$ would be sufficient to detect a medium effect size of $f^2 = 0.15$.

Measures and procedure. With the following exceptions, these were the same as in Study 1:

- A wrongness question was added: How wrong was Cynthia?
- The intention question was changed: To what extent did Cynthia intend to harm anyone?
- An outcome severity question was added: How severe (serious) was the outcome?
- The 'reckless' level of the negligence IV was removed
- The negligent vignette no longer included information about children often playing in leaf piles on the side of the road.
- The order of presentation of the judgment questions was randomized

Ethics and pre-registration. Ethical approval and pre-registration for Study 1 applied also to Study 2.

Data analysis. Data were analyzed as in Study 1 except that the added outcome severity was used instead of the outcome IV in the multiple regressions and mediation analyses. Preliminary analyses indicated no substantive main or interaction effects of the country of recruitment (UK or US), and so this factor was omitted from further analyses. For the multiple regressions, both perceived negligence and outcome severity were mean-centred. The maximum correlation (between perceived negligence and causal responsibility) was $r = .66$, and the maximum VIF was 1.91, indicating that multicollinearity was unlikely to have impacted the models. Except for

14 cases that were excluded from the punishment model, three from the blame model, and two from the wrongness model, the standardized residuals ranged between -3.29 and 3.29, indicating that the models were not biased by any individual cases.

Results

Punishment judgments. When no negligence information was given (the original condition), 48 out of 54 (89.9%) participants assigned no punishment when the outcome was positive, and none assigned more than a small fine; but when the outcome was negative, 54 out of 59 (91.5%) assigned punishment, of whom 41 (69.5%) thought the driver should be jailed for at least a year (Figure 3a).

As in Study 1, with few exceptions participants assigned high levels of punishment if, and only if, the agent was negligent *and* the outcome was negative. When the driver killed the children but was not negligent, 10 out of 57 (17.5%) participants assigned a jail sentence, and none did so when she was negligent but caused no harm. In contrast, of the 60 participants for whom both negligence and outcome were negative, 50 (83.3%) said she should be jailed, 17 (28.3%) for more than ten years.

An ANOVA indicated main effects of the outcome IV, $F(1,335) = 409.92, p < .001, \eta_p^2 = .55$, and of the negligence IV, $F(2,335) = 58.30, p < .001, \eta_p^2 = .26$: mean punishment ratings in the negligent and original conditions were significantly higher than when the driver was not negligent, $ps < .001$ (Figure 3a). There was also an interaction between outcome and negligence, $F(2, 335) = 50.41, p < .001, \eta_p^2 = .23$: when the outcome was positive, there was little or no negligence effect on punishment judgments, $F(2,162) = 2.82, p = .06, \eta^2 = .03$, whereas negligence strongly influenced punishment when the outcome was negative, $F(2,173) = 59.26, p < .001, \eta^2 = .41$.

Multiple regression indicated that punishment judgments were strongly influenced by outcome severity, perceived negligence, and the interaction between these factors. The effect of causal responsibility was relatively slight, and that of intention negligible (Table 8a).

Mediation analysis indicated that the total effect of outcome severity on punishment, $c = 0.50$, $p < .001$, 95% CI [0.44, 0.56] – meaning that on average participants assigned half a point more punishment for each 1-point increase on the 10-point perceived severity scale – was largely accounted for by its direct effect, $c' = 0.41$, $p < .001$, 95% CI [0.36, 0.46] (Figure 4). There was a modest indirect effect through perceived negligence, $a_1*b_1 = 0.08$, 95% CI [0.05, 0.11], but little or no evidence of mediation by intention, $a_2*b_2 = 0.01$, 95% CI [0.00, 0.02], or causation $a_3*b_3 = 0.01$, 95% CI [0.00, 0.01].

The total effect of perceived negligence on punishment, $c = 0.57$, $p < .001$, 95% CI [0.49, 0.65], resulted mainly from its direct effect, $c' = 0.36$, $p < .001$, 95% CI [0.28, 0.45]. There was an indirect effect of negligence through outcome severity, $a_4*b_4 = 0.14$, 95% CI [0.09, 0.19], but little or none through intention, $a_2*b_2 = 0.02$, 95% CI [0.00, 0.04] or causation, $a_3*b_3 = 0.05$, 95% CI [-0.01, 0.10]. The total effect of causation on punishment judgments, $c = 0.41$, $p = .004$, 95% CI [0.31, 0.51], was largely accounted for by its indirect effect through negligence, $a_1*b_1 = 0.26$, 95% CI [0.19, 0.33]. Its direct effect was not significant, $c' = 0.07$, $p = 0.18$, 95% CI [-0.01, 0.16].

Blame judgments. The high correlations (Table 7) and Figures 3b and 3d indicate that blame judgments were strongly associated with both the negligence IV and perceived negligence. ANOVA corroborated this substantial effect of the negligence IV, $F(2, 335) = 155.87$, $p < .001$, $\eta_p^2 = .48$. Blame ratings were similar when the driver was negligent and when no negligence information was given (the original condition), $p = .25$, but were higher at both these levels than when she was not negligent, $ps < .001$. The effect of outcome was modest, $F(1,$

335) = 23.11, $p < .001$, $\eta_p^2 = 0.07$, and there was no evidence of an interaction between negligence and outcome, $F(2, 335) = 1.30$, $p = .27$, $\eta_p^2 = .01$.

Multiple regression showed that both perceived negligence and, to a lesser extent, causal responsibility, predicted blame ratings (Table 8b). There was no indication that either outcome severity or intention predicted blame, nor that there was an interaction between negligence and outcome.

Mediation analysis indicated a modest total effect of outcome severity on blame judgments, $c = 0.18$, $p < .001$, 95% CI [0.10, 0.26], and that this was due primarily to an indirect effect through negligence, $a_1*b_1 = 0.14$, 95% CI [0.09, 0.20]. There was no evidence of outcome having a direct effect on blame, $c' = 0.00$, $p = .83$, 95% CI [-0.13, 0.05].

The total effect of perceived negligence on blame, $c = 0.86$, $p < .001$, 95% CI [0.80, 0.91] was primarily accounted for by its direct effect, $c' = 0.67$, $p < .001$, 95% CI [0.60, 0.75], but there was also an indirect effect through causation, $a_3*b_3 = 0.17$, 95% CI [0.12, 0.25].

The total effect of causal responsibility on blame, $c = 0.77$, $p < .001$, 95% CI [0.69, 0.85] partly resulted from its direct effect, $c' = 0.28$, $p < .001$, 95% CI [0.20, 0.35], but more from its indirect effect through negligence, $a_1*b_1 = 0.48$, 95% CI [0.40, 0.56].

Wrongness judgments. As with blame, wrongness judgments were strongly associated with both the negligence IV and perceived negligence, but less so with outcome (Table 7). ANOVA corroborated this strong effect of the negligence IV, $F(2, 335) = 107.52$, $p < .001$, $\eta_p^2 = .39$: wrongness judgments were slightly higher in the negligent than the original condition, $p = .001$, but in both were much higher than when the driver was not negligent, $ps < .001$ (Figure 3c). There was also a main effect of outcome, $F(1, 335) = 42.18$, $p < .001$, $\eta_p^2 = .11$, and a slight interaction with negligence, $F(2, 335) = 6.56$, $p = .002$, $\eta_p^2 = .04$.

Multiple regression indicated that wrongness judgments were strongly predicted by perceived negligence (Table 8c). These judgments were also predicted modestly, but

significantly, by ratings of outcome severity and of the driver's intention and causal responsibility.

Mediation analysis indicated that the moderate total effect of outcome severity on wrongness judgments, $c = 0.23$, $p < .001$, 95% CI [0.16, 0.31], was accounted for more by its indirect effect through negligence, $a_1*b_1 = 0.15$, 95% CI [0.10, 0.21], than by its direct effect, $c' = 0.06$, $p = .003$, 95% CI [0.02, 0.10]. In contrast, the total effect of perceived negligence on wrongness was substantial, $c = 0.81$, $p < .001$, 95% CI [0.76, 0.86], as was its direct effect, $c' = 0.71$, $p < .001$, 95% CI [0.65, 0.78]. There was also a small indirect effect of negligence through causal responsibility, $a_3*b_3 = 0.06$, 95% CI [0.01, 0.11].

The total effect of intention on wrongness, $c = 0.63$, $p < .001$, 95% CI [0.42, 0.84] was partly accounted for by its direct effect, $c' = 0.21$, $p < .001$, 95% CI [0.10, 0.32], but more so by its indirect effect through negligence, $a_1*b_1 = 0.36$, 95% CI [0.25, 0.49].

Perceived negligence. As in Study 1, perceived negligence was strongly correlated with the negligence IV, and more modestly with both the outcome IV and perceived severity of the outcome (Table 7). Correlations between perceived negligence and all three judgments were high, especially blame and wrongness.

Figure 3d shows the mean ratings of perceived negligence reported by the six groups. An ANOVA indicated that the negligence IV strongly influenced perceived negligence, $F(2, 335) = 213.85$, $p < .001$, $\eta_p^2 = .56$: perceived negligence was significantly lower in the non-negligent condition than in both the negligent and original conditions, $ps < .001$, and perceived negligence was slightly lower in the original than in the negligent conditions, $p = .05$.

There was also a significant main effect of outcome, $F(1, 335) = 46.42$, $p < .001$, $\eta_p^2 = .12$: the driver was considered more negligent when the children were killed. In addition, there was a slight interaction between the negligence IV and outcome, $F(2, 335) = 4.76$, $p = .01$, $\eta_p^2 = .03$: when the outcome was negative the influence of the negligence IV, $F(2, 173) = 151.27$, $p < .001$,

$\eta^2 = .64$, was stronger than when the outcome was positive, $F(2, 162) = 72.29, p < .001, \eta^2 = .47$.

Study 2 Discussion

The results of Study 2 are generally consistent with, and expand on, those of Study 1. In particular, the findings of substantial influences of negligence on punishment and blame, and also of outcome and the interaction between outcome and negligence on punishment but not blame judgments, were replicated. Causal responsibility also influenced blame judgments, though this was partially mediated by negligence.

Another main finding of Study 2 was that wrongness judgments are similar to blame judgments: they too are based primarily on negligence. This was consistent with the negligence-based account's seventh prediction. The two-process model's seventh prediction, according to which wrongness judgments would be based primarily on perceived intention, received limited support: intention had a substantial total effect on wrongness, but this was accounted for primarily by its indirect effect through negligence.

The outcome effects on perceived negligence and blame were again relatively slight, and this was also the case for wrongness. These effects on both blame and wrongness judgments were largely mediated by perceived negligence.

General discussion

In this research we investigated the relative influences of agents' negligence and their actions' unintended (accidental) outcomes on moral judgments. Of particular interest was whether negligence is a confounding variable that explains moral luck. In two studies we replicated Martin and Cushman's (2016) experiment in which a driver either unluckily killed two children, or luckily caused no harm, and we manipulated the driver's level of negligence. Participants assigned punishment, blame and wrongness, and also rated the driver's negligence, causal responsibility, and intentionality. Contrasting hypotheses from the two-process model

(Cushman, 2008; Cushman et al., 2013) and the negligence-based account of moral judgment were tested.

There were five main findings: The first was that, with few exceptions, participants assigned high levels of punishment when, and only when, the outcome was negative *and* the driver was negligent. Although we replicated the strong outcome effect reported in previous studies – running over the children was considered much more punishable than running over the sticks – for most participants this factor alone was not sufficient grounds for punishing the driver severely. In both studies, most punishment judgments followed a conjunction rule: if the driver was negligent *and* the outcome was negative, she deserved severe punishment; otherwise, she did not.

Second, blame judgments were based primarily on negligence. Causal responsibility also had a strong effect, although in both studies this was to a large extent mediated by negligence.

The third main finding was that wrongness judgments were influenced almost exclusively by negligence. Intention also had an effect, but this too was partly mediated by negligence.

Fourth, participants considered agents more negligent, blameworthy, punishable and wrong when the outcome was negative. However, in contrast to punishment judgments, the influence of outcome on blame and wrongness judgments was relatively slight and largely mediated by negligence.

And fifth, in the absence of negligence information (the original condition), participants tended to assume that the agent was approximately as negligent, punishable, blameworthy and wrong as in the negligent condition, in which her negligence was stated explicitly.

Regarding punishment judgments, our findings are similar to those of Martin and Cushman's (2016) original study. The current findings thus corroborate those of the original study, and show that they generalize to British samples.

The main predictions from the two-process model were that outcome – not negligence – would be the primary influence on punishment and blame judgments; that causal responsibility and intention would also have important roles; and that intention would have most influence on wrongness judgments. Regarding punishment judgments, there was some support for the model's predictions because there were, indeed, strong outcome effects: in particular, almost all participants assigned little or no punishment when the outcome was positive. However, the outcome effect was strongly moderated by negligence.

The findings were less consistent with the two-process model's predictions concerning blame judgments. In both studies the modest influence of outcome was largely accounted for by an indirect effect of negligence. Causal responsibility had some influence on these judgments, although this was partially mediated by negligence. Similarly, there was a modest effect of intention on wrongness judgments, but this too was partially mediated by negligence. The two-process model's prediction that causal responsibility would mediate negligence effects was not supported: in both studies the indirect effects of negligence on all three judgments through causation or intention were low or nil, while its direct effects were substantial. Perhaps most significantly, the two-process model did not predict the strong impact of negligence on punishment judgments, nor its substantial influence on blame and wrongness judgments.

In contrast, the negligence-based account predicted that negligence would be the principal influence on all three types of moral judgment. The findings reported here strongly support this prediction regarding blame and wrongness judgments, and partially support it regarding punishment judgments. Also consistent with this account, the outcome effects on blame and wrongness judgments were relatively slight and largely mediated by negligence. However, the negligence-based account did not predict the strong and independent influence of outcome on punishment judgments.

Together, Martin and Cushman's (2016) original study and the two reported here provide a clear example of moral luck: despite the driver having no control over whether the children were in the leaf pile, there was a strong outcome effect on punishment judgments. However, our findings shed new light on this intriguing phenomenon because they indicate two important qualifiers: first, an agent needs to be considered sufficiently negligent to be punished severely for an unlucky outcome; and second, moral luck applies much more to punishment than to blame and wrongness judgments.

Despite using different methods and analyses, two of our findings in particular are remarkably consistent with those of Kneer and colleagues (Kneer & Machery, 2019; Kneer & Skoczeń, 2021), neither of which had been reported previously. First, they too found stronger outcome effects on punishment than on blame and wrongness judgments; like us, they report that the influences on blame judgments seem to be more similar to those on wrongness. In fact, we found a slight direct outcome effect on wrongness judgments, but none at all on blame. This contrasts with the two-process model, according to which punishment and blame are influenced similarly by both intentions and outcomes, and it is only wrongness that is based almost exclusively on intentions and beliefs. Kneer and Machery (p. 182) suggest that these discrepant findings might reflect the phrasing of questions: whereas we and they asked about blameworthiness, Cushman (2008, p. 358) asked "How much blame does [agent] deserve?" Further research is required to test this and other possible explanations.

Second, Kneer and colleagues also found that perceived negligence plays a substantial role in moral judgments of accidental actions. Kneer and Machery (2019) reported that the outcome effects on wrongness and blame judgments were partly (Study 3) or wholly (Study 4a) mediated by negligence attributions. Kneer and Skoczeń (2021, Study 1) added a measure of subjective probability and found that this and negligence together accounted for the significant direct effect of outcome on blame (they did not include wrongness). Similarly, by manipulating its level and

measuring participants' perceptions, we showed the strong effect of negligence on punishment, blame and wrongness judgments, and that the slight outcome effects on blame and wrongness were largely mediated by negligence. In addition, we included causal responsibility and intention, which also influenced blame and wrongness judgments, respectively.

Both the present findings and Kneer and colleagues' are consistent with a form of dual process model in which blame and wrongness judgments of accidental actions are influenced by one process which is sensitive to negligence, and punishment judgments are influenced by this and a process which is sensitive to outcome. Our studies indicate that these two influences on punishment judgments are not additive but interact such that high levels of punishment are usually assigned only when both the outcome is negative and the agent is negligent.

Another intriguing and novel finding of both present studies was that negligence moderated the outcome effect on punishment judgments almost exclusively when the outcome was negative, that is, when the pattern of punishment judgments was similar to that of blame and wrongness judgments. It appears that, when the outcome is negative, punishment judgments are actually very similar to blame and wrongness judgments, particularly in that they are strongly influenced by negligence. However, when outcomes are positive, punishment judgments differ from blame and wrongness judgments; while blame and wrongness remain strongly influenced by negligence, assigned punishment becomes negligible or nil, regardless of the level of negligence.

When seen in this light it seems that, when outcomes are negative, there is little or no element of luck; agents are judged to be punishable to the extent that they are considered negligent and hence blameworthy and wrong. These punishment judgments are therefore based on rational, morally-relevant grounds, rather than on luck or outcome bias. In contrast, when outcomes are positive, punishment is not assigned even to agents who are judged to be negligent, blameworthy and wrong; these are the beneficiaries of irrational moral luck. In short, the

phenomenon of moral luck is largely about the *good* fortune of the negligent agent who is not punished because the outcome is positive, rather than about the *mis*fortune of the negligent agent who is punished because the outcome is negative.

The finding that participants considered the driver to be as negligent when they were given no negligence information (the original condition) as when they were told explicitly that she was negligent (the negligent condition) has several implications. First, it explains these participants' punishment judgments. Consistent with high levels of punishment being assigned when, and only when, the agent is considered negligent *and* the outcome is negative, participants in the original condition assumed that the driver was negligent, and so assigned little or no punishment when the outcome was positive, and high levels of punishment when the outcome was negative. Second, this may help explain the findings of other studies (e.g., Cushman, 2008; Gino et al., 2010; Kneer & Machery, 2019; Martin & Cushman, 2016; Robbennolt, 2000) in which outcome effects were reported when agents were not explicitly negligent: again, in the absence of this information, people assume negligence, and assign punishment accordingly. If this is the case, then outcome and negligence have often been confounded in previous research, with the result that outcome effects have been overestimated, and negligence effects underestimated. Future researchers are encouraged to avoid this confound and to test this possible explanation by both explicitly manipulating negligence and measuring perceived negligence.

Another important implication of this finding is that judgments can often be unjustly severe when no negligence information is provided. The results reported here indicate that, whenever judgments of accidental agents are made, for example in courts, workplaces, schools or homes, the advice should be to seek and provide this information if at all possible, even if it is that there was no evidence that the agent was negligent.

On the other hand, when the outcome is positive, there seems to be a strong and arguably unjust tendency to assign little or no punishment, irrespective of the action. The current findings

suggest that, in such circumstances, people making moral judgments should also be asked about negligence, blame and wrongness. If they say that the agent was negligent, blameworthy and wrong, but should not be punished, then this outcome-biased leniency should at least be challenged: to the rational judge who ignores outcomes because they are beyond the agent's control, the agent deserves punishment as much as someone who was equally negligent, but whose action unluckily led to a negative outcome.

For several reasons we urge caution in interpreting the findings of these experiments. First, we used only one set of vignettes, and it might be that others elicit different responses. However, Kneer and Machery (2019) and Kneer and Skoczeń (2021) both used two sets of vignettes, and several of their main findings are largely consistent with ours. The current evidence therefore points towards these findings generalizing across a range of stimuli.

Another possible criticism concerns causal direction, in particular whether perceived negligence influences judgments, or judgments influence perceived negligence. Alicke and colleagues (Alicke, 2014; Mazzocco et al., 2004) have argued that negligence ratings are post hoc rationalizations or validations of moral judgments that result from irrational outcome biases; that is, perceived negligence is based on a posteriori outcome information rather than on a priori negligence. In this sense, their model has much in common with other theories of moral judgment that also stress the role of intuitive processes as opposed to the rational application of moral principles (e.g., Greene & Haidt, 2002; Greene et al., 2008). However, in our experiments we both manipulated a priori negligence (the negligence IV, which could not be influenced by outcome), and measured perceived negligence (which could). That the a priori negligence IV strongly influenced all three types of moral judgment is, we think, beyond dispute, as is the finding that it was also the principal influence on perceived negligence. At least to this extent, then, perceived negligence cannot have been a post hoc rationalization of outcome-based moral judgments. If Alicke and colleagues were correct on this point, outcome would have strongly

predicted perceived negligence, but actually it did so only modestly. Instead, negligence is a substantial influence on judgments, independent of outcome.

The present studies benefited from systematically varying negligence as an IV with several levels – original (no negligence information); no negligence; negligence, and in Study 1, recklessness. As well as indicating the strong effect of negligence on moral judgments, this design enabled us to identify the outcome x negligence interaction (first main finding), to find that providing no negligence information leads participants to assume negligence (fifth main finding), and to show that moral luck seems to apply only to punishment judgments of agents whose actions luckily turn out well, rather than to unlucky agents whose actions turn out badly. The negligence IV also confirmed that it was negligence that influenced judgments, rather than judgments that influenced perceived negligence, a possible explanation of correlations between these factors that has been put forward by Alicke and colleagues (Alicke, 2014; Mazzocco et al., 2004). It served, too, to validate the measure of perceived negligence, in that participants consistently attributed much greater negligence when the driver was negligent than when she was not. And finally, the negligence IV increased variance in perceived negligence, and hence statistical power.

Although they did not manipulate negligence, a particular strength of Kneer and Machery's (2019) and Kneer and Skoczeń's (2021) studies is that they asked participants to rate the probability that the accident would occur. They found that participants considered the accidents more likely when the outcomes were negative than when they were neutral, that is, they ascribed probability post hoc. Moreover, they report that the outcome effects on negligence are mediated by these assessments of probability. Kneer and colleagues' analyses therefore add another level of explanation to ours because they shed light on the reasons why accidents are considered more negligent when outcomes are negative; specifically, it seems that in such circumstances people ascribe greater negligence because, given the negative outcome's assumed

higher probability, they think the agent should have been more aware of the greater risk of their action resulting in harm. Future researchers in this area are encouraged to include both these features – negligence as an IV as well as a perceived covariate, and a measure of probability assessment.

Closely related to probability is foreseeability (Lagnado & Channon, 2008; Margoni & Surian, 2021). We can speculate that individual differences in judgments – as opposed to variance in responses resulting from manipulation of the negligence and outcome IVs – reflect differences in views concerning how foreseeable it was that children were hiding in the leaf-pile. Participants who believed that this was simply unforeseeable, that is, Cynthia could not reasonably be expected to have known that children might be in the leaf-pile – would be expected to judge her blameless, whereas others who considered it more foreseeable – perhaps because they knew of children playing in leaf-piles – would judge her correspondingly more blameworthy. This possibility could be tested by asking participants how foreseeable they considered the outcome to be. In addition, as with Kneer and Machery's (2019) and Kneer and Skoczeń's measures of perceived probability, it would be interesting to assess the outcome effect on foreseeability by comparing participants' ratings following positive and negative outcomes.

The findings reported here – especially those concerning blame and wrongness judgments – are more consistent with the negligence-based account's predictions than with the two-process model's. However, they suggest the potential for a synthesis of the two in which negligence plays a central role in moral judgments of accidental harms. This could be accomplished through the role that beliefs play in the processing of others' mental states. Our results suggest that, when evaluating actions with unintended outcomes, people typically consider not only whether agents *had* beliefs about the potential consequences, but also whether they *should* have had these beliefs. The two-process model could thereby accommodate a substantial role for negligence in cases of unintended, though potentially foreseeable, accidents.

Otsuka (2009) has attempted to resolve the philosophical puzzle of moral luck and its contradiction of Kant's Control Principle by distinguishing between 'brute', or unavoidable moral luck, and 'option' moral luck, to which agents expose themselves through voluntary choices. Whereas the former is unfair because the agent had no choice, and the possible outcomes are entirely beyond their control, he argues that option moral luck is fair, and that it is therefore reasonable to blame and punish the unlucky agent. Otsuka draws an analogy with gambling, and points out that we do not consider it unfair when one roulette-player wins a fortune, while another loses one, both entirely by luck; after all, both players chose to expose themselves equally to the same risks of winning and losing.

Of course, the victim of bad option luck is condemned not because they *intend* to cause harm, any more than the losing gambler *intends* to lose a fortune. But if it is not malicious intention that leads us to condemn the victim of bad moral luck, what is it? Otsuka proposes that:

“... what gives rise to resentment and indignation [i.e., blame] is the gratuitousness of bad behaviour [...] when the person indulges in it *even though he knows*, or *ought to have known*, that he could have behaved decently instead and it was reasonable to expect him to have done so.” (Otsuka, 2009, p. 329; emphases added)

According to Otsuka, then, it is reasonable to blame victims of bad option moral luck because they are *reckless*, that is, they knowingly take a risk, or *negligent*, because they ought to have known the risk they were taking.

Cynthia provides a good example of option moral luck because, when she chose to drive through the leaf-pile, she knowingly (i.e., recklessly), or unknowingly (i.e., negligently), exposed herself to the risk of causing harm. In effect, she took a gamble, and so, according to Otsuka, it is fair for her to be assigned blame and punishment when the gamble does not pay off, that is, when her option moral luck is bad and she kills the children. Similarly, when Cynthia

unavoidably drove through the leaf-pile to prevent a car accident with the same result, she was the victim of bad brute luck and so it would be unfair to condemn her.

To an extent, our findings are consistent with and corroborate Otsuka's analysis; participants considered it fair and appropriate to punish Cynthia when she chose to drive through the leaf-pile and killed the children (she was the victim of bad option moral luck), but not to punish her when she did not kill them (she was the beneficiary of good option moral luck). Moreover, they condemned her not because of the morally-irrelevant outcome per se (which the two-process model would predict), but because of her morally-relevant recklessness or negligence. And, also in line with Otsuka's argument, participants did not condemn Cynthia when she couldn't avoid driving through the leaf-pile and so, when the children were killed, was the victim of bad brute luck.

But our findings contrast with Otsuka's analysis in that, like most other moral philosophers (e.g., Hartman, 2017; Nagel, 1979; Nelkin 2019, Otsuka, 2009; Williams 1981), he focuses on blame when discussing moral luck. Like Kneer and Machery's (2019), our evidence indicates that moral luck has much more to do with punishment than with blame. This might reflect a difference between ethicists' beliefs and lay people's intuitions, or perhaps these philosophers are simply mistaken to assume that moral luck applies primarily to blame. Our findings indicate that people attribute blame and wrongness according to how negligent they perceive the agent to be, and that they do so pretty much regardless of the outcome. Moreover, even regarding punishment judgments, there seems to be only good, not bad, moral luck. In short, it seems that lay people intuitively apply the Control Principle except when they make punishment judgments of option (not brute) good (not bad) luck. If this is the case, then the problem of moral luck is much more specific and focused than moral philosophers and psychologists have typically realised.

In summary, the main proposal of the negligence-based account – that negligence exerts a strong, direct influence on punishment, blame and wrongness judgments of accidental agents – was supported by the current findings. Negligence was the principal influence on blame and wrongness, and punishment judgments were based on both negligence and outcome: most participants assigned severe punishment when, and only when, the outcome was negative *and* the agent was negligent.

It was also proposed that moral luck occurs because accidental agents are considered more negligent, and therefore more punishable, blameworthy and wrong, when their actions have negative rather than positive outcomes. Regarding blame and wrongness judgments, this differential attribution of negligence to lucky and unlucky agents was indeed found to account for the relatively slight outcome effects. However, it did not explain the much greater influence of outcome on punishment, nor why moral luck seems primarily to apply to punishment judgments when outcomes are positive.

These findings also indicate that, when participants are not given information to the contrary (as has typically been the case in previous studies of moral judgment), many assume that accidental agents are negligent, and therefore blameworthy, wrong and – when the outcome is negative – deserving of punishment. If this is the case, the implication is that judgments that appear to be based on outcome (which should be morally-irrelevant because the accidental agent lacks control over it) are often actually based on the morally-relevant factor of perceived negligence, and that, as a result, the influence of outcome on moral judgments has often been overestimated in previous research, and that of negligence underestimated.

References

- Alicke, M. D. (2014). Evaluating blame hypotheses. *Psychological Inquiry*, 25(2), 187-192.
<https://doi.org/10.1080/1047840X.2014.902723>
- Costanzo, P. R., Coie, J. D., Grumet, J. F., & Farnill, D. (1973). A reexamination of the effects of intent and consequence on children's moral judgments. *Child Development*, 44, 154-161. <https://www.jstor.org/stable/1127693>
- 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., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, 127(1), 6-21. <https://doi.org/10.1016/j.cognition.2012.11.008>
- Dunfield, K. A., & Kuhlmeier, V. A. (2010). Intention-mediated selective helping in infancy. *Psychological Science*, 21(4), 523-527. <https://doi.org/10.1177/0956797610364119>
- Enzle, M., & Hawkins, W. (1992). Prior negligence mediates a posteriori outcome effects on moral judgment. *Journal of Experimental Social Psychology*, 28, 169-185.
[https://doi.org/10.1016/0022-1031\(92\)90037-K](https://doi.org/10.1016/0022-1031(92)90037-K)
- Farnill, D. (1974). The effects of social-judgment set on children's use of intent information. *Journal of Personality*, 42, 276-289. <https://doi.org/10.1111/j.1467-6494.1974.tb00674.x>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191. <https://doi.org/10.3758/BF03193146>
- Gino, F., Moore, D. A., & Bazerman, M. H. (2009). *No harm, no foul: The outcome bias in ethical judgments* (Harvard Business School working paper, Issue).
https://www.hbs.edu/faculty/Publication%20Files/08-080_1751f2c7-abe2-402b-9959-1d8190ebf62a.pdf
- Gino, F., Shu, L. L., & Bazerman, M. H. (2010). Nameless + harmless = blameless: When seemingly irrelevant factors influence judgment of (un)ethical behavior. *Organizational Behavior and Human Decision Processes*, 111(2), 93-101.
<https://doi.org/10.1016/j.obhdp.2009.11.001>
- Greene, J., & Haidt, J. (2002). How (and where) does moral judgment work? *Trends in Cognitive Sciences*, 6, 517-523. [https://doi.org/10.1016/S1364-6613\(02\)02011-9](https://doi.org/10.1016/S1364-6613(02)02011-9)

- Greene, J., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*, 1144–1154.
[https://doi.org/10.1016/S1364-6613\(02\)02011-9](https://doi.org/10.1016/S1364-6613(02)02011-9)
- Hamlin, J. K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, *128*, 451–474.
<http://dx.doi.org/10.1016/j.cognition.2013.04.004>
- Hartman, R. (2017). *In defense of moral luck: Why luck often affects praiseworthiness and blameworthiness*. Routledge. <https://doi.org/113829344X>
- Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2 ed.). Guilford Press.
- Killen, M., Mulvey, K. L., Richardson, C., Jampol, N., & Woodward, A. (2011). The accidental transgressor: Morally-relevant theory of mind. *Cognition*, *119*(2), 197–215.
<https://doi.org/10.1016/j.cognition.2011.01.006>
- Kneer, M., & Machery, E. (2019). No luck for moral luck. *Cognition*, *182*, 331–348.
<https://doi.org/10.1016/j.cognition.2018.09.003>
- Kneer, M., & Skoczén, I. (2021). *Outcome effects, moral luck and the hindsight bias*.
<https://philpapers.org/rec/KNEOEM>
- Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, *108*, 754–770.
<https://doi.org/10.1016/j.cognition.2008.06.009>
- Laurent, S. M., Nuñez, N. L., & Schweitzer, K. A. (2016). Unintended, but still blameworthy: the roles of awareness, desire, and anger in negligence, restitution, and punishment. *Cognition and Emotion*, *30*(7), 1271–1288.
<https://doi.org/10.1080/02699931.2015.1058242>
- Lench, H. C., Darren Domsky, Smallman, R., & Darbor, K. E. (2015). Beliefs in moral luck: When and why blame hinges on luck. *British Journal of Psychology*, *106*, 272–287.
<https://doi.org/10.1111/bjop.12072>
- Margoni, F., Geipel, J., Hadjichristidis, C., & Surian, L. (2019). The influence of agents' negligence in shaping younger and older adults' moral judgment. *Cognitive Development*, *49*, 116–126. <https://doi.org/10.1016/j.cogdev.2018.12.002>
- Margoni, F., & Surian, L. (2016). Explaining the U-shaped development of intent-based moral judgments. *Frontiers in Psychology*, *7*. <https://doi.org/10.3389/fpsyg.2016.00219>

- Margoni, F., & Surian, L. (2020). Conceptual continuity in the development of intent-based moral judgment. *Journal of Experimental Child Psychology, 194*, 1-12.
<https://doi.org/10.1016/j.jecp.2020.104812>
- Margoni, F., & Surian, L. (2021). Judging accidental harm: Due care and foreseeability of side effects. *Current Psychology*.
- Martin, J. W., & Cushman, F. (2016). The adaptive logic of moral luck. In J. Sytsma & W. Buckwalter (Eds.), *The Blackwell companion to experimental philosophy* (pp. 190-202). John Wiley & Sons.
- Mazzocco, P. J., Alicke, M. D., & Davis, T. L. (2004). On the robustness of outcome bias: No constraint by prior culpability. *Basic and Applied Social Psychology, 26*, 131-146.
<https://doi.org/10.1080/01973533.2004.9646401>
- Mulvey, K. L., Gönültaş, S., & Richardson, C. B. (2020). Who Is to blame? Children's and adults' moral judgments regarding victim and transgressor negligence. *Cognitive Science, 44*(4). <https://doi.org/10.1111/cogs.12833>
- Nagel, T. (1979). *Mortal questions*. Cambridge University Press.
- Nelkin, D. K. (2021). Moral luck. *The Stanford Encyclopedia of Philosophy*
<https://plato.stanford.edu/archives/sum2021/entries/moral-luck/>
- Nobes, G., Panagiotaki, G., & Engelhardt, P. (2017). The development of intention-based morality: The influence of intention salience and recency, negligence, and outcome on children's and adults' judgments. *Developmental Psychology, 53*, 1895-1911.
<https://doi.org/10.1037/dev0000380>
- Nobes, G., Panagiotaki, G., & Pawson, C. (2009). The influence of negligence, intention and outcome on children's moral judgments. *Journal of Experimental Child Psychology, 104*, 382-397. <https://doi.org/10.1016/j.jecp.2009.08.001>
- Núñez, N. L., Laurent, S. M., & Gray, J. M. (2014). Is negligence a first cousin to intentionality? Lay conceptions of negligence and its relationship to intentionality. *Applied Cognitive Psychology, 28*, 55–65. <https://doi.org/10.1002/acp.2957>
- Otsuka, M. (2009). Moral luck: Optional, not brute. *Philosophical Perspectives, 23*, 373-388
<https://www.jstor.org/stable/40658409>
- Patil, I., Calò, M., Fornasier, F., Young, L., & Silani, G. (2017). Neuroanatomical correlates of forgiving unintentional harms. *Scientific Reports, 7*, 1-10.
<https://doi.org/10.1038/srep45967>
- Piaget, J. (1932). *The moral judgment of the child*. Routledge and Kegan Paul.

- Rizzo, M. T., Li, L., Burkholder, A. R., & Killen, M. (2019). Lying, negligence, or lack of knowledge? Children's intention-based moral reasoning about resource claims. *Developmental Psychology, 55*, 274–285. <https://doi.org/10.1037/dev0000635>
- Robbennolt, J. K. (2000). Outcome severity and judgments of “responsibility”: A meta-analytic review. *Journal of Applied Social Psychology, 30*, 2575-2609. <https://doi.org/10.1111/j.1559-1816.2000.tb02451.x>
- Shen, F. X., Hoffman, M. B., Jones, O. D., Greene, J. D., & Marois, R. (2011). Sorting guilty minds. *New York University Law Review, 86*, 1306-1360.
- Shultz, T. R., & Wright, K. (1985). Concepts of negligence and intention in the assignment of moral responsibility. *Canadian Journal of Behavioural Science, 17*, 97–108. <http://dx.doi.org.uea.idm.oclc.org/10.1037/h0080138>
- Van de Vondervoort, J. W., & Hamlin, J. K. (2018). Preschoolers focus on others' intentions when forming sociomoral judgments. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2018.01851>
- Walster, E. (1966). Assignment of responsibility for an accident. *Journal of Personality and Social Psychology, 3*, 13–19. <https://doi.org/10.1037/h0022733>
- Williams, B. (1981). Moral luck. In *Moral luck* (pp. 20-39). Cambridge University Press.
- Young, L., Nichols, S., & Saxe, R. (2010). Investigating the neural and cognitive basis of moral luck: It's not what you do but what you know. *Review of Philosophy and Psychology, 1*, 333–349. <https://doi.org/10.1007/s13164-010-0027-y>
- Zelazo, P. D., Helwig, C. C., & Lau, A. (1996). Intention, act, and outcome in behavioral prediction and moral judgment. *Child Development, 67*, 2478–2492. <https://doi.org/10.1111/j.1467-8624.1996.tb01869.x>

Table 1

Vignette Texts, Varied by Negligence Level and Outcome

Stem	One autumn afternoon, Cynthia is driving home from getting lunch. Her two brothers are in the car with her. As they get close to their home, they see a very large leaf pile on the side of the road, the biggest they have seen all season.
Negligence level	
<i>Original (no negligence information)</i>	Cynthia's younger brother asks her to drive through the leaf pile, and she agrees. Cynthia drives the car through the leaf pile, and she feels two bumps as the car passes over some objects.
<i>Non-negligent¹</i>	Suddenly a car making an illegal turn swerves wildly towards them, and to prevent a crash Cynthia unavoidably steers the car through the leaf pile, despite her best efforts not to. As she drives through the leaf pile, she feels two bumps as the car passes over some objects.
<i>Negligent</i>	Cynthia's younger brother asks her to drive through the leaf pile. Wanting to please her brother, Cynthia agrees and drives the car through the leaf pile, forgetting the fact that in this area children often play in leaf piles on the side of the road. As she drives through the leaf pile, she feels two bumps as the car passes over some objects.
<i>Reckless²</i>	Cynthia's younger brother asks her to drive through the leaf pile. Wanting to please her brother, Cynthia agrees and drives the car through the leaf pile, even though she knows that in this area children often play in leaf piles on the side of the road. As she drives through the leaf pile, she feels two bumps as the car passes over some objects.
Outcome	
<i>Positive³</i>	Unknown to Cynthia, there were some large sticks in the leaf pile. The car is not damaged by the sticks.
<i>Negative</i>	Unknown to Cynthia, two children were hiding in the leaf pile. The children were hit by the car and killed instantly.

¹ To avoid confusion with the blame DV we use the term 'non-negligent' for Shen et al.'s (2011) 'blameless'.

² For brevity, and in line with the common perception that the terms 'negligent' and 'reckless' are often interchangeable (Shen et al., 2011) – and our findings here are consistent with this view – recklessness is referred to here as an extreme form of negligence.

³ 'Positive outcome' is used in its relative, rather than absolute sense: running over sticks is actually a neutral or non-negative outcome

Table 2

Hypotheses concerning influences on punishment and blame judgments (and wrongness judgments, Study 2 only) from a) the Two-Process Model (Cushman et al., 2008; 2013); and b) the Negligence-Based Account¹

	Two-process model	Negligence-based account
H1. Outcome effects on judgments	Outcome effects will be substantial and direct: judgments will be low when the outcome is positive (no harm), and relatively high when negative (children killed), though not <i>very</i> high because the driver did not intend to kill the children.	The influence of outcome will be slight, although the driver will be judged more leniently when the outcome is positive (see H3).
H2. Negligence effects on judgments	Negligence effects will be slight or nil, although the driver might be judged more severely when she is reckless (see H4).	Negligence effects will be substantial: judgments will be low when the driver is not negligent, moderate when negligent, and relatively high when reckless. In the original condition they will be moderate, i.e., similar to when the driver is negligent.
H3. Mediation of outcome effects on judgments	Outcome effects will not be mediated by negligence	The slight outcome effect (see H1) will be largely or wholly mediated by perceived negligence
H4. Mediation of negligence effects on judgments	Any negligence effects (see H2) will be mediated by perceived intention and / or causal responsibility	Negligence effects will be largely or wholly direct, i.e., unmediated
H5. Influences of perceived intention and causal responsibility on judgments	Judgments will be relatively severe when the agent is considered causally responsible for, and / or to have intended, the negative outcome	

H6. Differences between judgments There will be no substantial differences in the influences on punishment and blame judgments; these judgments will therefore be moderately or highly correlated

H7. Wrongness judgments (Study 2)	Wrongness judgments will be primarily influenced by intention; any outcome or negligence effects on wrongness will be mediated by perceived intention	Wrongness judgments will be primarily influenced by negligence, for the same reasons and in the same ways as punishment and blame judgments
-----------------------------------	---	---

¹Predictions set out in the preregistration concerning a “near-miss” effect were inadequately tested in Study 1, and were not tested in Study 2, and so are not considered here.

Table 3

Questions and response rating scales

Question	Response
1. How much should Cynthia be punished?	0 – No punishment at all; 1 = Community service; 2 = A small fine and some probation; 3 = A large fine or a few days in jail; 4 = A few weeks in jail; 5 = a few months in jail; 6 = A year in jail; 7 = 2-3 years in jail; 8 = 4-7 years in jail; 9 = 7-10 years in jail, 10 = More than 10 years in jail
2. Why did you make the response that you did?	Open text
3. To what extent was Cynthia blameworthy for the way things turned out?	0 = Not at all blameworthy; 5 = Somewhat blameworthy; 10 = Extremely blameworthy
4. To what extent was Cynthia negligent (careless or reckless)?	0 – Not at all negligent – as careful as possible; 5 – Somewhat negligent – could have been more careful; 10 – Extremely negligent – should have been much more careful
5. To what extent did Cynthia cause things to turn out the way they did?	0 – Didn't cause it at all; 5 – Somewhat; 10 – Caused it entirely
6. To what extent did Cynthia intend things to turn out the way they did? (Study 1)	0 – Not at all (100% unintentional); 5 – Somewhat intentional; 10 – Completely (100% intentional)
6. To what extent did Cynthia intend to harm anyone? (Study 2)	0 = No intention at all to harm anyone; 5 = Moderate intention to harm; 10 = Fully intended to harm someone
Study 2 only	
7. How wrong was Cynthia?	0 = Not at all wrong; 5 = Moderately wrong; 10 = Extremely wrong
8. How severe (serious) was the outcome?	0 = Not at all serious or severe; 5 = Moderately serious or severe; 10 = Extremely serious or severe

Table 4

Means, SDs and Pearson Correlation Coefficients between IVs, Judgments and Perceived Covariates (N = 343)

	IVs		Judgments		Perceived				
	<i>M</i>	<i>SD</i>	Outcome	Negligence ¹	Punishment	Blame	Negligence	Causation	Intention
Outcome IV			-						
Negligence IV ¹			.05	-					
Punishment	3.09	3.75	.72**	.31**	-				
Blame	4.67	3.62	.13*	.56**	.43**	-			
Negligence	5.14	3.65	.20**	.62**	.54**	.78**	-		
Causation	5.37	3.48	.12*	.49**	.35**	.72**	.69**	-	
Intention	1.60	2.58	-.41**	.23**	-.21**	.28**	.22**	.27**	-

¹ *n* = 261 (original / no negligence information group not included)

* *p* < .05. ** *p* < .01

Table 5

Linear Regression Models: Predictors of a) Punishment Judgments and b) Blame Judgments (Experiment 1)

	Punishment						Blame					
	<i>b</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	95% CI	<i>b</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	95% CI
Constant	2.67	.23		11.37	.000	2.20, 3.15	2.73	.27		10.08	.000	2.06, 3.38
Outcome (sticks or children)	4.60	.22	.62	20.64	.000	4.15, 5.06	-0.02	.26	.00	-0.06	.954	-0.61, 0.57
Perceived negligence	0.42	.04	.41	11.08	.000	0.34, 0.50	0.54	.04	.55	12.44	.000	0.43, 0.65
Causation	0.02	.04	.02	0.52	.60	-0.05, 0.10	0.34	.05	.33	7.63	.000	0.22, 0.46
Intention	0.02	.04	.01	0.44	.66	-0.05, 0.10	0.09	.05	.06	1.64	.102	-0.05, 0.23
Outcome*negligence	0.63	.06	.30	11.24	.000	0.52, 0.73	-0.03	.07	-.02	-0.50	.619	-0.17, 0.10
Maximum Cook's distance				.08						.07		
Standardized residuals range				-3.29, 3.28						-3.20, 2.82		
R^2				.78						.69		

Table 6

Characteristics of Participants, by Country of Recruitment

	UK	US	Total
<i>n</i>	168	173	341
Age (years)			
Mean	38.3	39.8	39.1
SD	22.4	10.8	17.5
Range	18-94	22-72	18-94
Gender			
% female	76.5	41.2	58.6
Nationality			
% British	83.0	0.0	40.9
% European	8.5	0.0	4.2
% US	0.6	100.0	51.0
First language			
% English	88.6	99.4	94.1
Education			
% Post-18	89.7	82.2	86.9

Table 7

Means, SDs, and Pearson Correlation Coefficients between IVs, Judgments and Perceived Covariates (N=341)

			IVs		Judgments			Perceived		
	<i>M</i>	<i>SD</i>	Outcome	Negligence ¹	Punish	Blame	Wrong	Negligence	Cause	Intention
Outcome IV			-							
Negligence IV ¹			-.01	-						
Punishment	2.04	3.28	.65**	.41**	-					
Blame	4.32	3.63	.18**	.76**	.57**	-				
Wrongness	3.72	3.34	.26**	.70**	.64**	.81**	-			
Negligence	4.63	3.59	.24**	.79**	.61**	.84**	.86**	-		
Causation	5.43	2.96	.06	.68**	.39**	.70**	.62**	.66**	-	
Intention	0.22	0.63	.07	.18**	.23**	.20**	.27**	.20**	.18**	-
Outcome severity	5.27	4.69	.97**	.03	.66**	.21**	.30**	.27**	.10	.11*

¹ $n = 228$ (original / no negligence information group not included)

* $p < .05$. ** $p < .01$.

Table 8

Linear Regression Models: Predictors of a) Punishment, b) Blame, and c) Wrongness Judgments

	Punishment						Blame						Wrongness					
	<i>b</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	95% CI	<i>b</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	95% CI	<i>b</i>	<i>SE</i>	β	<i>t</i>	<i>p</i>	95% CI
Constant	1.85	.16		11.55	.000	1.54, 2.17	3.09	.24		12.86	.000	2.50, 3.68	3.65	.33		16.76	.000	3.19, 4.08
Outcome severity	0.42	.01	.56	28.92	.000	0.39, 0.45	0.03	.02	.00	0.12	.904	-0.04, 0.05	0.05	.02	.07	2.53	.011	0.01, 0.09
Perceived negligence	0.43	.03	.45	17.59	.000	0.39, 0.49	0.71	.04	.70	19.04	.000	0.61, 0.79	0.73	.03	.78	21.80	.000	0.66, 0.80
Causation	0.05	.03	.05	2.00	.047	.001, 0.101	0.26	.04	.24	6.77	.000	0.16, 0.36	0.08	.04	.08	2.22	.027	0.01, 0.15
Intention	0.09	.07	.02	1.28	.202	-0.48, 0.23	0.06	.11	.02	0.59	.559	-0.08, 0.31	0.32	.10	.09	3.26	.001	0.13, 0.58
Outcome*Negligence	0.09	.00	.41	22.27	.000	0.08, 0.10	0.00	.01	.01	0.19	.853	-0.01, 0.01	0.01	.01	.06	2.38	.002	0.03, 0.02
Maximum Cook's distance				.04						.11						.14		
Standardized residuals range				-3.29, 3.19						-3.27, 3.24						-2.91, 3.11		
R^2				.89						.77						.78		

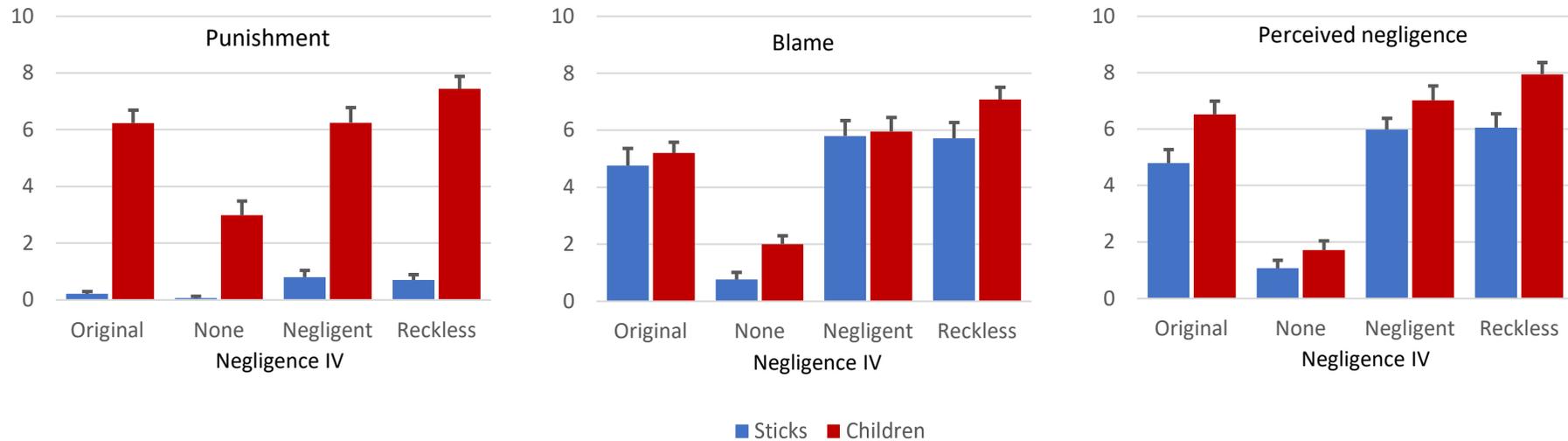


Figure 1. Mean (+ SE) ratings of a) punishment b) blame, and c) perceived negligence (0 = not at all / none; 10 = very high / > 10 years in jail) by the manipulated negligence IV (original = no negligence information; none = non-negligent) and accident outcome (blue = sticks broken; red = children killed).

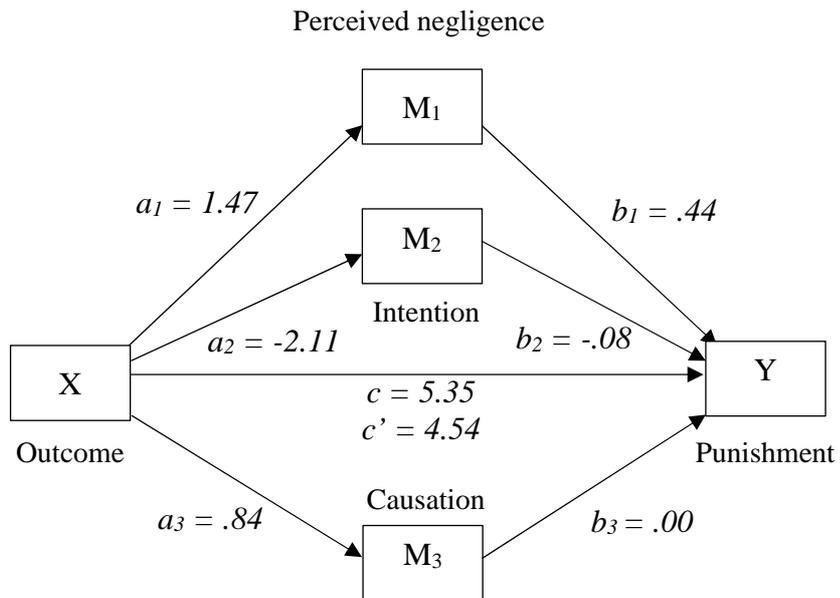


Figure 2. Mediation model: The substantial outcome effect on punishment, c , was largely accounted for by its direct effect, c' .

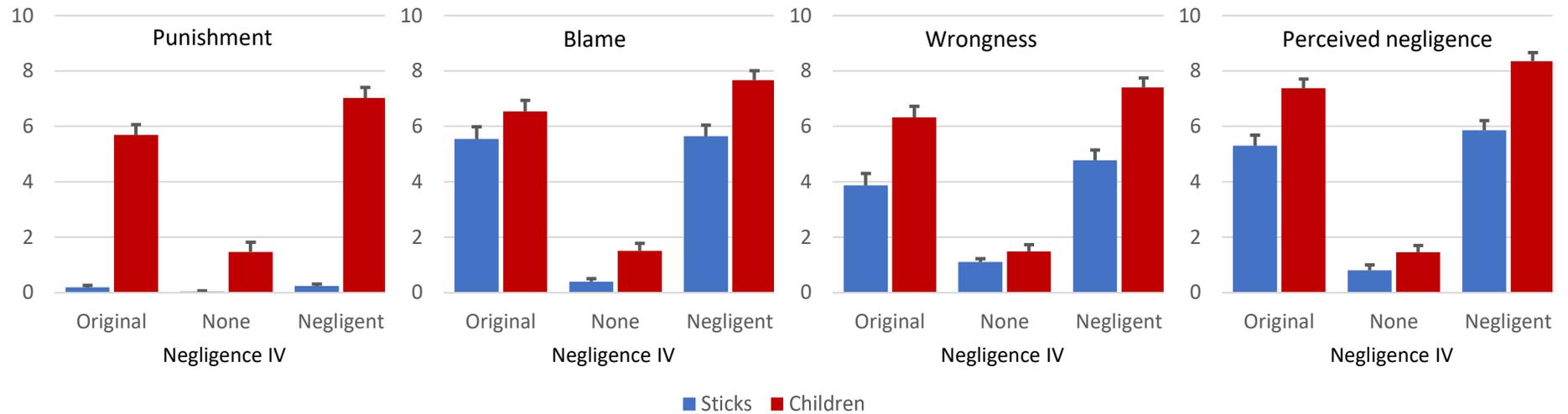


Figure 3. Mean (+ SE) ratings of a) punishment, b) blame, c) wrongness, and d) perceived negligence (0 = not at all / none; 10 = very high / > 10 years in jail) by the manipulated negligence IV (original = no negligence information; none = non-negligent) and accident outcome (blue = sticks broken; red = children killed).

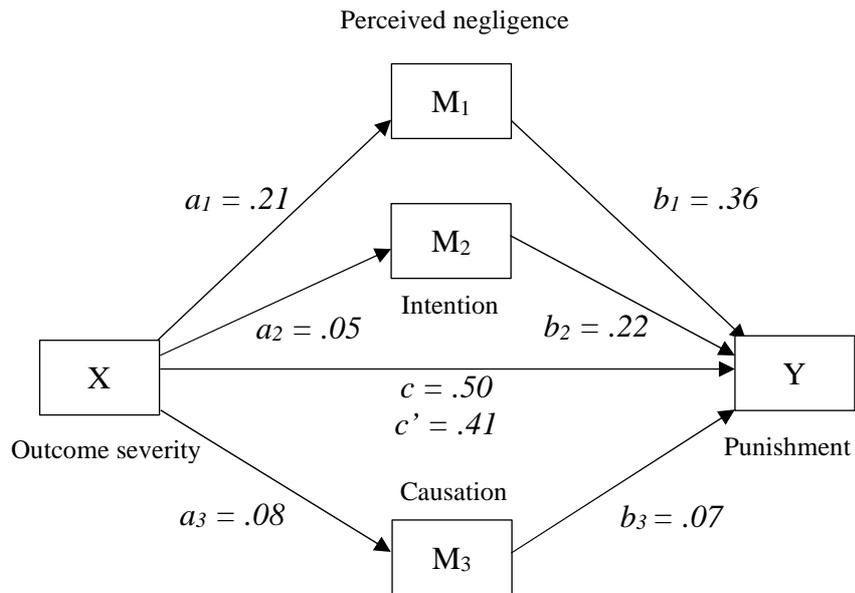


Figure 4. Mediation model: The total effect of outcome severity on punishment, c , was accounted for primarily by its direct effect, c' , rather than by any indirect effect through negligence, intention or causal responsibility