Moral luck and the roles of outcome and negligence in moral judgments ☆

Gavin Nobes [a],[*], Georgia Panagiotaki [b], Justin W. Martin [c]

[a] *School of Psychology, University of East Anglia, Norwich, UK*
[b] *Norwich Medical School, University of East Anglia, Norwich, UK*
[c] *Department of Psychology and Neuroscience, Boston College, MA, United States of America*

A R T I C L E   I N F O

A B S T R A C T

Two experiments were conducted to investigate the influences of outcome and negligence on moral judgments of accidental actions, and hence their roles in the explanation of moral luck. In Experiment 1 ($N = 300$), two previous studies were replicated in which an agent armed with either a bat or a gun (to manipulate negligence) unintentionally killed a suspected intruder who turned out, luckily, to be a burglar, or unluckily, a family friend (to manipulate outcome). In response to an online questionnaire, participants made moral judgments of punishment, blame and wrongness and rated the agent's negligence and intentionality. The effects of both outcome (victim) and negligence (weapon type) IVs were slight, whereas *perceived* negligence had a substantial impact on all three judgments. In Experiment 2 ($N = 241$) the potential influence of both outcome and negligence was raised by increasing the contrasts between conditions: the agent was armed or unarmed, and the suspected intruder was harmed or unharmed. Perceived negligence again had a substantial impact on all three judgments, but now outcome, too, had a strong and direct effect on punishment judgments. These findings indicate that outcome effects on blame and wrongness judgments of accidental agents result primarily from the differential attribution of negligence: agents are considered more negligent – and hence more culpable – when outcomes are worse. In contrast, high levels of punishment are usually assigned when, and only when, the accidental agent is considered negligent *and* the outcome is negative. We discuss the implications for the interpretation of previous findings of strong outcome effects, and whether these effects, and therefore moral luck, are best explained by hindsight bias or by more rational updating of moral judgments.

On April 16, 1984, Peter Garmess came home early from work and heard noises upstairs. He knew it was not a member of his family because his wife and daughter were away. He went upstairs, got his gun, and opened his daughter's door. Seeing a shadow move, Garmess had no time to think, and he fired his gun. His bullet hit Lawrence Drake, killing him instantly. It turned out that Garmess was correct in his assumption that Drake was a burglar. Drake had spent time in prison on three different occasions for robbery and had been responsible for a number of burglaries in the area.

When Alicke and Davis (1989) asked participants to judge Garmess, their mean blame and punishment ratings were relatively low on the 0–10 scales (3.61 and 2.53, respectively). However, when the same vignette ended with Drake being an innocent family friend, Garmess was judged to deserve significantly more blame and punishment (5.78 and 4.53, respectively). Since all other aspects of the vignettes were

identical, this and many related studies (e.g., Bornstein, 1998; Gino, Shu, & Bazerman, 2010; Kurdi, Krosch, & Ferguson, 2020; Lench, Domsky, Smallman, & Darbor, 2015; Martin & Cushman, 2016; McNamara, Willard, Norenzayan, & Henrich, 2019; Patil, Calò, Fornasier, Young, & Silani, 2017; Robbennolt, 2000; Shen, Hoffman, Jones, Greene, & Marois, 2011; Walster, 1966; Young, Nichols, & Saxe, 2010) show that outcomes strongly affect people's moral judgments.

The outcome effect is puzzling because agents are considered culpable and punishable for reasons – in this case, Drake's identity – that are entirely beyond their control (Kant, 1784/1998). The Garmess vignettes therefore provide a good example of moral luck (Hartman, 2017; Nagel, 1979; Nelkin, 2021; Williams, 1981) because, despite his intentions and actions being the same, Garmess was judged leniently or severely depending on whether his victim luckily turned out to be a dangerous intruder, or unluckily, an innocent family friend. As Mazzocco, Alicke, and Davis (2004) suggest, "This tendency is difficult to

---

reconcile with any standards of judgmental rationality or reasonableness" (p. 145). The intriguing phenomenon of moral luck appears to be consistent with theories that stress the role of intuitive and emotional processes in judgments, as opposed to the rational application of moral principles (e.g., Alicke, 2000; Greene, 2008; Haidt, 2001).

The large majority of researchers in this area have focused on the influence of intention and outcome, "the two major building blocks of moral judgment" (Kurdi et al, 2020, p. 2). Typically they have manipulated intention and outcome using a 2 × 2 design ("The workhorse of this literature", Cushman, 2015, p. 97) and report that both directly and independently influence moral judgments. But there are other factors. One such is negligence which, with some exceptions (e.g., Kneer & Machery, 2019; Kneer & Skoczeń, 2023; Laurent, Nuñez, & Schweitzer, 2016; Margoni, Geipel, Hadjichristidis, & Surian, 2019; Margoni & Surian, 2021; Nobes & Martin, 2022; Nobes, Panagiotaki, & Engelhardt, 2017; Shultz & Wright, 1985; Siegal & Peterson, 1998), has not been manipulated, measured or controlled for in previous studies. As a result, much less is known about the influence of negligence than of intention and outcome on moral judgments, particularly of accidental agents, and this is reflected in recent models of moral judgment – some of which are discussed below – that tend to pay little attention to its influence.

The negligence-based account (Nobes & Martin, 2022; Nobes, Panagiotaki, & Pawson, 2009; see also Kneer & Machery, 2019) has been developed to address this relatively poor understanding of its role in moral judgment. There are three main proposals. First, especially when actions are unintentional, people's attributions of culpability are strongly influenced by their perceptions of the agents' negligence. Consider, for example, two drivers who share the same benign motive – to get home without causing any harm – but who both accidentally cause the death of a pedestrian. Despite the identical intentions and outcomes, we judge them very differently: the one who drove carefully and took all reasonable precautions walks free, whereas the other who drove without due care and attention – perhaps by speeding and using their phone – is likely to be punished severely. We do not claim that this is an original or surprising insight. What is surprising is how often researchers in this field have overlooked the influence of negligence, especially given that it exerts such a strong effect on legal decisions and everyday moral judgments.

The second main proposal of the negligence-based account is that, when no information is given about an accidental harm-doer's level of negligence, people tend to assume that they must have been negligent. Nobes and Martin (2022) found this to be the case: participants who were not told whether an agent who caused an accident was cautious or careless considered her to be as negligent as did those who were told that she was negligent (see also Margoni et al., 2019).

The third is the differential attribution of negligence, that is, the tendency to assume that worse outcomes of accidents were caused by greater negligence. This might result from a hindsight bias (e.g., Bernstein, Erdfelder, Meltzoff, Peria, & Loftus, 2011; Hawkins & Hastie, 1990; Kneer & Skoczeń, 2023), and in the General Discussion we put forward an alternative explanation based on updating of assessments of negligence, and hence of culpability, as outcome information is provided.

This account does more than add negligence to the usual intention / outcome dichotomy as a likely third factor in the explanation of moral judgment (proposal 1): it also provides a possible explanation of the intriguing outcome effect on moral judgments, and hence of moral luck. This is because, in the large majority of studies that report strong outcome effects, no negligence information has been provided. This means that participants' judgments of accidental harm-doers might not be based on the harmful outcome per se (a "pure" outcome effect), but instead on the assumption that these agents were negligent (proposal 2). Since in such studies negligence has not been manipulated, and both outcome and perceived negligence covary with judgments, it is not possible to know whether the judgments are driven by outcome, negligence, or both. If the negligence-based account is correct, moral luck

occurs not because of outcome per se (i.e., whether an agent unluckily causes harm or luckily doesn't), but because we differentially attribute negligence according to these outcomes (proposal 3). (See also Kneer & Machery, 2019, and Nobes & Martin, 2022.)

To illustrate these points, an example of the many studies in which moral judgment has been investigated by manipulating intention and outcome, but not negligence, is Cushman (2008). In one set of vignettes, Jenny took a sculpture class, and either did, or did not, burn her partner's hand while welding. When she did not intend any harm, the participants who were told that the outcome was negative assigned her substantially more blame and punishment (though not wrongness or permissibility) than those for whom the outcome was neutral. Cushman's interpretation is that outcome independently and strongly influenced judgments. From a negligence-based perspective, however, this outcome effect arises not from the outcome per se, but from participants perceiving Jenny to be more negligent, and therefore culpable, when she burnt her partner's hand. Since they were not told whether Jenny took any precautions to prevent the accident, it seems reasonable of participants to assume that she didn't, and to consider her negligent for not doing so; they might think, for instance, that she should have checked with the teacher whether her partner's hand could be burnt, or asked her partner to wear welding gloves. If correct, this explanation illustrates the negligence-based account's first proposal, because their perceptions of negligence strongly influenced the participants' judgments; and its second proposal, because, in the absence of information to the contrary, they thought she should have taken care not to burn her partner, but didn't; and its third proposal, because participants attributed negligence differentially depending on whether or not she burnt her partner's hand. According to this account, therefore, these participants' judgments covaried with outcome and so gave the impression that they were outcome-based, but they were actually based on perceived negligence.

Two further examples are Barrett et al.'s (2016) and McNamara et al.'s (2019) cross-cultural studies of moral judgments. In their accidental physical harm vignette, participants were told that two unacquainted men were at a crowded party. The victim was sitting down when the agent "tripped and fell forward and his hand accidentally struck [the victim] in the face, severely bruising his eye" (Barrett et al., Supplementary Information, pp. 53–54). The researchers reported wide variation in the extent to which intentions and outcomes influenced moral judgments in different cultures. For instance, McNamara et al. report that the Yasawan people of Fiji judged the negative outcomes of accidental harms more severely than the ill intentions of attempted harms, from which the researchers infer that these people's judgments are based more on outcomes than on intentions. However, such findings are also consistent with (and, we would argue, better explained by) Yasawans being strongly influenced by intentions, but also being particularly sensitive to negligence, such that someone who accidentally but carelessly causes harm – by, for example, not looking where he was going and clumsily tripping on to and injuring a seated party guest – is considered at least as culpable as someone who causes it deliberately.

Since in studies like these examples – and there are many more – negligence has not been manipulated or measured, we cannot be sure whether it is outcome or negligence, or both, that is the principal driver of judgments of accidental actions. If negligence had been manipulated or measured in these studies, the researchers might have found that apparently outcome-based judgments were actually based on the perceived negligence of the agents, in which case the results would have been interpreted very differently. The implication is that findings concerning the outcome effect might need to be radically reinterpreted, and models of moral judgment that don't include negligence might need to be restructured.

In the experiments reported here we used Alicke and Davis' (1989) Garmess vignettes to investigate these issues concerning the explanation of moral judgments of accidental agents, the relative influence of outcome and negligence on these judgments, and whether negligence might account for the outcome effects, and hence moral luck. We did so

by manipulating both outcome and negligence experimentally so that the influence of different levels of each could be directly observed, and by measuring participants' perceptions of negligence, to assess how these were related to outcomes and judgments. We tested hypotheses from the negligence-based account, and from four other views, or models, of moral judgment.

A rare example of a study in which negligence was manipulated was conducted by Enzle and Hawkins (1992). They replicated Alicke and Davis' (1989) study, and again found an outcome effect on blame judgments (they did not ask about punishment) when Garmess used a gun. However, they also varied negligence by telling half their participants that he used a baseball bat instead of a gun. This manipulation was validated by participants in two studies rating the ownership and use of guns to be more negligent than of bats. Enzle and Hawkins found that when Garmess was armed with a bat there was no outcome effect; he was considered equally blameworthy whether he killed a burglar or innocent friend (Fig. 1a). They therefore proposed that the outcome effect reported by Alicke and Davis occurred only because Garmess was considered highly negligent to have and to use a gun. According to Enzle and Hawkins, there was an interaction between negligence and outcome such that Garmess was judged to be very blameworthy only when he acted with substantial negligence (he used a gun) *and* the outcome was very negative (the apparent intruder was actually an innocent friend).

Mazzocco et al. (2004) tested this hypothesis in a series of studies. When they replicated Enzle and Hawkins' experiment they too found that judgments of blame were increased if the agent used a gun rather than a bat (Fig. 1b). This was also the case for perceived negligence (again validating this manipulation) and punishment. However, they reported that the outcome effect occurred regardless of weapon type; that is, they did not find the interaction between outcome and negligence that Enzle and Hawkins reported. In their Study 4, Mazzocco et al. then further reduced Garmess' negligence by saying that he was unarmed, and merely called out, so that he could not possibly have foreseen Drake's death in the way that he perhaps should have done when he armed himself with a gun. Even in this condition of very low negligence they again found that agents were blamed slightly but significantly more when his victim was innocent than when he was a burglar. Mazzocco et al. therefore disagreed with Enzle and Hawkins, and argued that outcome effects persist even when negligence is very low.

Alicke and colleagues explain these findings in terms of their culpable control model (CCM; Alicke, 2000, 2014; Mazzocco et al., 2004). The CCM is a "blame-early" (Malle, Guglielmo, & Monroe, 2014) or "motivated-blame" account (Monroe & Malle, 2019), according to which we have an outcome bias that leads us to base our judgments of agents primarily and spontaneously on the outcomes of their actions. They propose that inferences about causes and mental states such as negligence are actually post hoc justifications, or validations, of these judgments. Other motivated-blame models that share with the CCM the view that intuitive or emotional moral judgments precede, or "anchor",

and bias processing of mental and causal information include those of Haidt (2001), Greene (2008), and Pettit and Knobe (2009); for further discussion see Monroe and Malle (2019).

Cushman and colleagues (Cushman, 2008; Cushman, Sheketoff, Wharton, & Carey, 2013; Martin & Cushman, 2016) have proposed the two-process model of moral judgment, according to which wrongness judgments – to do with right and wrong, good and bad, acceptability and impermissibility – are influenced by a process that is sensitive to mental states, in particular intentions and beliefs. In contrast, blame and punishment judgments are influenced by both this and a second process, which is sensitive to the causes of outcomes. This means that agents who unintentionally cause harm – such as when Garmess shot the innocent friend – are not considered wrong (because intentions are absent), but they are considered moderately blameworthy and punishable (because they cause, though do not intend, the outcomes). This, they argue, explains the phenomenon of moral luck, and also why it applies primarily or exclusively to blame and punishment, rather than to wrongness judgments. At least regarding blame and punishment, the two-process model resembles Alicke and colleagues' CCM account in that outcome directly influences moral judgments of accidental agents.

In contrast, according to Malle and colleagues' (Malle et al., 2014; Monroe & Malle, 2019) Path Model of blame, accidental agents are judged to be blameworthy if, as well as having caused harm, they *should* have prevented it (obligation), and they *could* have prevented it (capacity). The latter is "both the cognitive capacity to foresee the event and the physical capacity to actually prevent it… [where] the cognitive capacity to prevent… is the basis for the legal concept of negligence" (Malle et al., 2014, p. 155). According to this model, blame judgments are the products not of an outcome bias (cf. Alicke and colleagues), nor of a causally-sensitive process (cf. Cushman and colleagues) but of systematic information processing that is influenced by morally-relevant factors, including negligence.

For two reasons, Malle and colleagues play down the role of outcome in moral judgments. First, they point out that outcome effects are often modest; in a meta-analysis Robbennolt (2000) reported average correlations between outcome severity and responsibility of .08, and between outcome severity and blame of .17. Indeed, in Mazzocco et al.'s replication of Enzle and Hawkins' Garmess experiment, the outcome effect sizes (η^2) on punishment was .11, and on blame, only .07 (i.e., they explained 11% and 7% of the variance, respectively).

Second, Malle et al. propose that "outcome effects are… often readily explained by causal and mental state inferences mediating the outcome-blame relationship." (p. 162). In particular, since harm is much more likely to result from an ill-intentioned or negligent action than from a well-intentioned or non-negligent one, it is reasonable to assume that negative outcomes tend to indicate negative intentions or negligence and therefore culpability, and positive outcomes tend to indicate positive intentions or non-negligence and therefore innocence.

The negligence-based account builds on these explanations of the outcome effect in moral judgments of unintentional actions. As in the CCM and the two-process model, it acknowledges that people's judgments of accidents are often strongly influenced by outcomes. However, like Enzle and Hawkins (1992), it focuses on the role of negligence, and, like Malle et al.'s (2014) Path Model, it emphasizes the relative rationality, or systematicity, of these judgments, and how they are based on morally-relevant information such as negligence, capacity and obligation. Instead of a direct, unmediated outcome-judgment link that, according to Alicke (2000, 2014), results from an outcome bias, or, according to Cushman (2008), from a causally-sensitive process, the outcome effect on judgments is primarily or solely indirect and mediated by negligence.

The negligence-based account is largely consistent with the Path Model of blame, but it differs from, or expands on it, in three ways. First, rather than outcome exerting only a minor influence on moral judgments, it recognizes, and attempts to explain, outcome's often substantial – though, crucially, indirect – impact; second, the negligence-
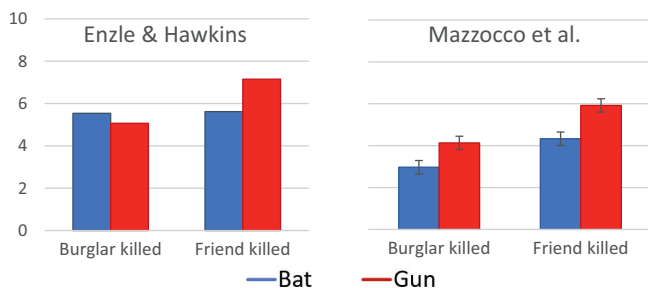

Fig. 1. Mean ratings [+ *SEs*] of Garmess' blameworthiness (maximum = 10) reported by a) Enzle and Hawkins (1992; Study 2, *N* = 104) and b) Mazzocco et al. (2004; Study 1, *N* = 219), by outcome (burglar or friend killed) and weapon type (to manipulate negligence). SEs not reported by Enzle and Hawkins.

Table 1
Predictions of relative effect sizes of factors on three moral judgments of garmess according to five accounts of moral judgments of accidental actions.

Factor Judgment	Account or model				
	Enzle and Hawkins (1992)	Culpable control model (Mazzocco et al., 2004)	Two-process model (Cushman, 2008)	Path Model (Malle et al., 2014)	Negligence-based (Nobes & Martin, 2022)
Outcome					
Punishment	–	High	Moderate	–	Moderate
Blame	Moderate	High	Moderate	Nil	Nil / low, mediated by negligence
Wrongness			Nil	–	
Negligence					
Punishment	–	Nil		–	Moderate
Blame	Moderate	Nil	Nil / low, mediated by intention	High	High
Wrongness				–	High
Outcome x negligence interaction					
Punishment	–	Nil	Nil	–	High
Blame	Moderate	Nil	Nil	Nil	Nil
Wrongness				–	Nil
Intention					
Punishment	–		Moderate	–	Nil / low
Blame	–		Moderate	Moderate	Nil / low
Wrongness			High	–	Moderate
Causal responsibility					
Punishment	–	Moderate	High	–	Low
Blame	–	Moderate	High	Moderate	Low
Wrongness			Nil	–	Low

based account extends the Path Model's focus on blame to punishment and wrongness judgments; and third, this account positions outcome as a major source of the information on which judgments are based and updated (see the General Discussion).

In a recent study, Nobes and Martin (2022) tested the negligence-based account using vignettes developed by Martin and Cushman (2016). In the two original stories, a driver steered her car through a leaf-pile, and outcome was manipulated by having either some sticks, or two children, hidden in the leaf-pile. As in the original study, there was a strong outcome effect on punishment judgments, $\eta_p^2 = .52$: when the driver ran over sticks, 3.1% assigned a jail sentence, whereas when she ran over the children, 79.8% thought she should be jailed for several months or more. Nobes and Martin also manipulated negligence by describing the driver as reckless, negligent or non negligent. Their findings largely supported the main points of the negligence-based account, in particular the key role played by negligence in blame and wrongness judgments of unintentional harms, and the mediation of the outcome effect on these judgments by perceived negligence. However, an unexpected finding was that, although punishment judgments, too, were strongly influenced by negligence, there was also a substantial and direct (i.e., unmediated) outcome effect; the large majority of participants assigned high levels of punishment – though not blame or wrongness – when, and only when, *both* the outcome was negative *and* the driver was negligent.

Similarly, Kneer and colleagues (Kneer & Machery, 2019; Kneer & Skoczeń, 2023) report that negligence has a substantial impact on punishment, blame and wrongness judgments, and that there was a strong direct outcome effect on punishment, but not on blame or wrongness. Also consistent with the negligence-based account, their participants considered actions to be more negligent when consequences were negative than when they were neutral, and this differential attribution of negligence accounted for the outcome effect on blame and wrongness judgments.

Both Kneer and colleagues' and Nobes and Martin's (2022) findings therefore support aspects of the two-process model, especially in that wrongness was influenced primarily by a mental state (in these cases negligence rather than intention, a distinction that we discuss in the General Discussion), and punishment by this same mental state but also by the outcome. In contrast to the two-process model, though, these researchers found that blame judgments patterned with wrongness rather than punishment; as with wrongness, the outcome effect on blame was relatively low, and largely mediated by negligence.

The present research: We sought to test these different accounts of the influences on moral judgments of accidental harms, and the reasons for the outcome effect – and hence of moral luck – by replicating Enzle and Hawkins' (1992) and Mazzocco et al.,'s (2004) experiments using the same Garmess vignettes. Our key research questions were: What are the relative influences of outcome and negligence on moral judgments?; and, Are moral judgments influenced by outcome per se, or are outcome effects partly or wholly explained by the differential attribution of negligence?

Participants were asked to judge Garmess' deserved punishment, blameworthiness and wrongness. In addition, they rated his negligence, and the extent to which he intended the outcome. In this way we built on the Path Model's, CCM's, and Enzle and Hawkins' focus on blame judgments, and Mazzocco et al.'s study, who asked about punishment, blame and negligence. Wrongness judgments and intention ratings were also included, in particular to compare findings with those of Nobes and Martin (2022), and to test the prediction from Cushman and colleagues' two-process model that, compared with punishment and blame, wrongness judgments would be influenced more by the extent to which actions are considered intentional, and correspondingly less by the severity of the outcome.

We tested predictions from the five accounts discussed above (Table 1). Enzle and Hawkins' (1992) findings indicate that perceived negligence should vary with weapon type ("a priori negligence")

because having and using a gun should be considered more negligent than having and using a baseball bat. In addition, greater negligence should lead participants to deem Garmess more blameworthy. However, these findings indicate that there should be an interaction in that the outcome – killing a friend rather than a burglar – will be considered more blameworthy only when negligence is particularly high, i.e., when Garmess used a gun.

According to Alicke and colleagues' CCM (Alicke, 2000, 2014; Mazzocco et al., 2004), Garmess should be judged more harshly when he kills the friend rather than the burglar. In line with Mazzocco et al.'s findings, using the gun rather than the bat should also be considered more culpable and to deserve more punishment, but no interaction between weapon type and outcome would be predicted. Perceived negligence should be associated with these judgments, but only because negligence ratings are post hoc rationalizations of the judgments, rather than perceived negligence being an influence on them.

From the perspective of the two-process model (Cushman, 2008; Cushman et al., 2013), the causally-sensitive process should lead participants to judge him blameworthy and punishable because Garmess caused the deaths. And, since killing a friend is a worse outcome than killing a burglar, this should be assigned more blame and punishment, i. e., there should be an independent outcome effect on these judgments. In contrast, wrongness judgments should not be directly affected by outcome because they are sensitive only to mental states. Since it does not include negligence as a factor, the two-process model either makes no predictions about its influence on judgments, or it predicts that negligence has little or no effect.

According to Malle et al.'s (2014) Path Model, the key influences on participants' judgments of Garmess' blameworthiness should be their perceptions of whether Garmess *should* and *could* have prevented Drake's death, that is, whether he had the obligation and the capacity to do so. Participants might be expected to think that Garmess had an obligation to prevent the death of a friend more than that of a burglar, but since he had no way of knowing (i.e., he lacked the cognitive capacity to foresee) his victim's identity, the outcome should not make a difference. However, he did have the capacity to choose whether to own and use a gun or a bat, and guns are usually considered more dangerous than bats. The Path Model should therefore predict an effect of weapon type – with use of a gun being more blameworthy than use of a bat – but not an outcome effect, i.e., the identity of the victim should not influence judgments. Since, as Malle et al. (2014, p. 155) point out, capacity and negligence are closely related concepts, perceived negligence should also predict the extent to which Garmess is deemed to be blameworthy.

The negligence-based account leads to the prediction that negligence will be the principal influence on judgments of Garmess' unintentional killings (proposal 1). Participants will assume that Garmess was negligent because, for example, he could have called out to the apparent intruder to find out who he was, or phoned the police, but didn't. Moreover, they will deem him more negligent when he arms himself with a gun than with a bat because, as Enzle and Hawkins (1992) and Mazzocco et al. (2004) reported, owning and using guns is considered more negligent than owning and using bats (proposal 2). There will also be an outcome effect, but this will be explained by differing perceptions of negligence (proposal 3). This is because participants will perceive Garmess to be more negligent when his victim, Drake, is a family friend than when he is a burglar, and they will judge him according to this perceived negligence. Judgments will therefore co-vary with outcome – and so give the impression of being influenced by it – but it is the negligence indicated by the outcome, not the outcome per se, that is the moral basis of the judgments. Specifically, the outcome-judgment path should be mediated by perceived negligence. This prediction is similar to that deriving from the Path Model's, that perceptions of negligence drive moral judgments, but extends it backwards to explain outcome effects, too.

Although not predicted by the negligence-based account, the findings of the Nobes and Martin (2022), Kneer and Machery (2019) and

Kneer and Skoczeń (2023) studies lead also to the expectation of a difference between punishment judgments on the one hand, and blame and wrongness judgments on the other: while all three will be strongly influenced by perceived negligence, there should also be a direct outcome effect on punishment, and an interaction such that high levels of punishment are assigned when, and only when, Garmess is considered negligent *and* he killed the friend.

Our focus in this research was on unintentional actions (accidental harms), but it is possible that some participants interpreted Garmess' actions as intentional; for example, killing with a gun could be considered more intended (or, at least, less *un*intended) than killing with a bat because guns are more lethal, or because owning a gun in the first place indicates a willingness to use and possibly kill with it. On the other hand, killing with a bat might be thought more intended because wielding one requires more effort than pulling a trigger. There is also strong evidence of intention attributions being influenced by outcomes (e.g., Kneer & Bourgeois-Gironde, 2017; Knobe, 2003; Rosset, 2008). To this extent, the two-process model would predict that intention had a strong influence on wrongness, and a moderate influence on punishment and blame judgments. The Path Model recognizes the fundamental role of intention regarding blame, while Nobes and Martin (2022) reported that it had some influence on wrongness judgments of the driver, but little or none on punishment or blame.

The Garmess vignettes used in previous research were phrased in terms of what Garmess *claimed* he did, rather than what he *actually* did. Some participants might have doubted this testimony and suspected instead that Garmess' version was an attempt to justify and excuse his actions. The result would be that their judgments were not wholly based on factors that are relevant to our research questions, notably outcome and negligence. We tested this possibility by giving half the participants the original, "testimonial", versions of the vignettes, and the other half adapted, "factual", accounts of what happened.

1. Experiment 1

1.1. Method[1]

1.1.1. Design

The relative influence of negligence and outcome information on moral judgments of accidental agents was investigated using a 2 (weapon type [bat, gun]) x 2 (victim identity [burglar, friend]) x 2 (vignette version [original, factual]) between-subjects design. Weapon type was varied to manipulate negligence (owning and using a gun is assumed to be more negligent than owning and using a bat), and victim identity to manipulate outcome (killing a burglar is a more positive outcome than killing an innocent friend). The original version was that used by Enzle and Hawkins (1992) and Mazzocco et al. (2004) in which the vignette's events are presented as Garmess' own testimony, whereas in the factual version they are stated as facts.

The dependent variables were the three moral judgments – deserved punishment, blame, and wrongness – and the two covariates were perceived negligence and intention. All five of these variables were measured on 0–10 Likert scales.

1.1.2. Recruitment and sample

A total of 404 people responded to invitations distributed on social media, and by email to a university psychology student participant panel. Of these, 72 (17.8%) did not start the questionnaire or provided no relevant data, and 21 (5.2%) reported not properly understanding the vignette and/or questions. Another four (1.0%) were excluded because they asked for their data not to be used. Seven (1.7%) more were excluded because two or all three of their judgments were outliers that

---

[1] See the Supplementary Material for further information on power, data analysis, and vignettes

indicated misunderstanding or lack of engagement; for example, saying that killing Drake the family friend was not at all punishable, blameworthy or wrong.

Sensitivity analyses (Perugini, Gallucci, & Costantini, 2018) using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) for the MANOVA (8 groups, 3 DVs) and linear multiple regressions (4 predictors) indicated that the sample of 300 participants was sufficient to detect even small effect sizes (f^2s = .026 and .063, [η_p^2s = .025 and .059] respectively) with α = .05 and power = .95.

The participants were all 18 or over, their mean age was 28.6 years, and 212 (70.7%) were women. The large majority – 244 (81.3%) – had continued education at least until 18. 250 (83.3%) described themselves as British, 15 (5%) European, and 14 (4.7%) American. The first language of 268 (87.3%) was English; there were also 8 (2.7%) native Cantonese or Mandarin speakers, and 5 (1.7%) native Spanish speakers.

1.1.3. Measures

Vignette stem (factual version)

On April 16, 1984, Peter Garmess *struck / shot* and fatally wounded Lawrence Drake. Garmess *struck Drake with a baseball bat / shot Drake with a gun* when he came home from work early and found Drake in his daughter's bedroom. Upon entering his house Garmess heard loud noises being made upstairs. He knew that it was not a member of his family because his wife and daughter were spending the week at their beach house. He decided that he better take the *baseball bat that he kept under his bed with him / gun that he kept in his bedroom drawer with him (the gun was properly licensed)*. He went very quietly up the stairs, got the *bat / gun*, and opened the door to his daughter's bedroom. At this point he saw a shadow move to the left inside the doorway. There was no time to think things out. He quickly rounded the corner of the door and *swung the bat / fired the gun*. His *blow / bullet* hit Drake, killing him instantly.

Positive outcome. As it turned out, Garmess was correct in his assumption that Drake was a burglar. Drake had spent time in prison on three different occasions for robbery and had been responsible for a number of burglaries in the area.

Negative outcome. As it turned out, Drake was the boyfriend of Garmess' daughter, and his daughter had left Drake keys to the house so that Drake could bring some of her things to the beach house, where he would be spending the day. Garmess had not met Drake and did not know what he looked like.

Questions. The first three questions and scales were those asked by Mazzocco et al. (2004), except that we added the words in italics to clarify the terms:

1. To what extent is Garmess blameworthy for his actions? (0 - Not at all blameworthy; 10 - Extremely blameworthy).
2. To what extent is Garmess negligent *(careless or reckless)* for his actions? (0 - Not at all negligent; 10 - Extremely negligent).
3. How severely would you sentence *(punish)* Garmess for his actions? (0 - No sentence at all; 10 = Very severe sentence).
   The fourth and fifth questions were new:
4. How right or wrong was Garmess to act as he did? (0 – Not wrong at all; 10 = Very wrong).
5. Before Garmess opened the door to his daughter's room, to what extent did he intend to kill the person inside? (0 = Didn't intend to kill at all; 10 = Fully intended to kill the person).

1.1.4. Procedure

The invitations to take part were linked to the online Qualtrics questionnaire. This began with information about the study and a consent form, and then participants answered demographic questions. The Garmess scenario was one of four that each participant responded to (the other three concerned other aspects of morality and are not reported here). The order of scenarios and of questions about Garmess, and the version of the Garmess vignette (according to the outcome, weapon type and testimonial or factual version) presented to each participant were allocated randomly and equally. Following completion of the Garmess vignette, and again at the end of the whole questionnaire, participants were asked whether they felt that they had understood the stories and questions, and whether they wished us to keep their data.

1.1.5. Ethics

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

1.1.6. Disclosures

We report all measures, manipulations, and exclusions in these studies.

1.1.7. Data analysis

First, the influence of the IVs was tested by conducting a 2 (Weapon type [bat, gun]) x 2 (Outcome [burglar, friend]) x 2 (Version [original, factual]) between-subjects ANOVA with perceived negligence as the DV, and then an equivalent MANOVA with punishment, blame and wrongness judgments as the DVs.

Second, the influence of perceived negligence (as opposed to the weapon type IV), outcome and intention was investigated by running multiple regressions with these variables, and the perceived negligence-outcome interaction, as predictors of the three judgments.

Third, mediation analyses were conducted to investigate the possible indirect effect of outcome on the three judgments through perceived negligence and intention.[2] Hayes' (2017) PROCESS method uses ordinary least squares path analysis, and calculates 95% confidence intervals based on 5000 bootstrap samples.

1.2. Results[3]

1.2.1. The influence of the manipulated IVs (negligence, outcome and vignette version) on perceived negligence and judgment ratings

A 2 (weapon type: [gun, bat]) x 2 (outcome [burglar, friend]) x 2 (vignette version [original, factual]) between subjects ANOVA with perceived negligence as the DV indicated an outcome effect – participants considered killing a criminal less negligent than killing an innocent friend, $F(1, 292) = 68.26$, $p < .001$, $\eta_p^2 = .19$ – and a modest weapon (negligence IV) effect – using the gun was perceived to be more negligent than using the bat, $F(1, 292) = 11.66$, $p = .001$, $\eta_p^2 = .04$ (Fig. 2a). There was also an interaction between outcome and weapon type, $F(1, 292) = 7.53$, $p = .006$, $\eta_p^2 = .03$: when Garmess used a bat, the outcome effect, $F(1,138) = 56.07$, $p < .001$, $\eta_p^2 = .29$, was stronger than when he used a gun, $F(1,158) = 16.32$, $p < .001$, $\eta_p^2 = .09$ (Fig. 2a). Neither the main effect of vignette version, $F(1, 292) = 1.37$, $p = .242$, $\eta_p^2 = .01$, nor its interactions with outcome, $F(1, 292) = 0.05$, $p = .832$, $\eta_p^2 < .01$, or weapon type, $F(1, 292) = 1.39$, $p = .240$, $\eta_p^2 = .01$, approached significance.

A MANOVA with the same IVs and the three judgments – punishment, blame and wrongness – as the DVs indicated an outcome effect on judgments, $F(3,290) = 20.91$, $p < .001$, $\eta_p^2 = .18$: killing the innocent friend was judged more severely than killing the burglar. There was also a main effect of weapon type, $F(3, 290) = 7.74$, $p < .001$, $\eta_p^2 = .07$, and a modest outcome x weapon interaction, $F(3, 290) = 3.75$, $p = .011$, $\eta_p^2 = 0.04$. This pattern was consistent across judgment types; the clearest difference was that the outcome effect on punishment, $F(1, 292) = 53.17$, $p < .001$, $\eta_p^2 = .15$, was stronger than on blame, $F(1, 292) = 12.82$,

---

[2] In the General Discussion we discuss the possibility that judgments might be causally antecedent to perceived negligence, in which case judgments mediate outcome effects on perceived negligence, rather than vice versa (cf. Alicke, 2000, 2014; Mazzocco et al., 2004). See also the Supplementary Material for reverse causal path analyses.

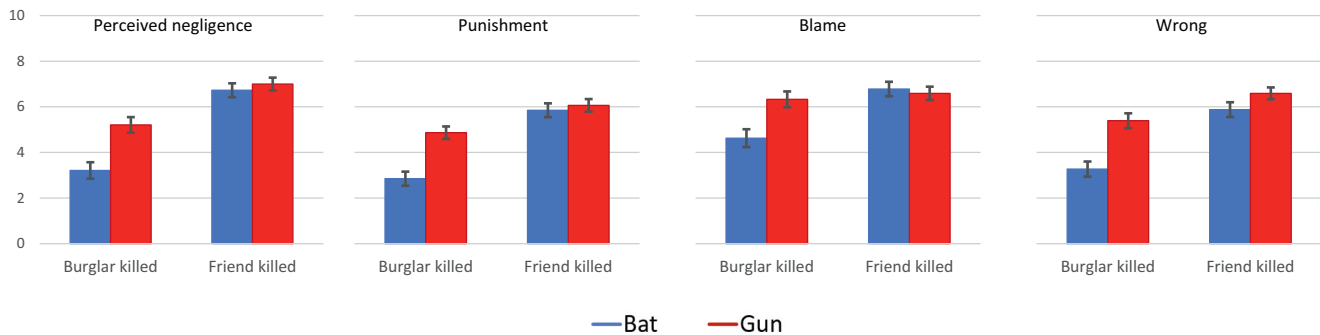[3] See the Supplementary Material for correlations and path analyses

Fig. 2. Mean [+ *SE*] ratings of a) perceived negligence, b) punishment, c) blame, and d) wrongness (0 = not at all / none; 10 = very high) by outcome (burglar or friend killed) and weapon type (bat or gun). (Experiment 1.)

$p < .001$, $\eta_p^2 = .04$ (Figs. 2b-d).

The outcome x weapon interactions were explored further to compare findings with those of Enzle and Hawkins (1992) and Mazzocco et al. (2004). When Garmess used a bat there was a clear outcome effect: when he killed the friend rather than the burglar, participants judged him more punishable, $F(1,138) = 47.75$, $p < .001$, $\eta_p^2 = .26$, blame-worthy, $F(1,138) = 18.61$, $p < .001$, $\eta_p^2 = .12$, and wrong, $F(1,138) = 31.38$, $p < .001$, $\eta_p^2 = .19$. In contrast, when he used a gun, outcome had less impact: Garmess was considered only slightly more punishable, $F(1,158) = 9.50$, $p = .002$, $\eta_p^2 = .06$, and wrong, $F(1,158) = 8.18$, $p = .005$, $\eta_p^2 = .05$, and no more blameworthy, $F(1,158) = 0.33$, $p = .568$, $\eta_p^2 = .00$, when he shot the friend. Although there was no main effect of vignette version, $F(3, 290) = 0.76$, $p = .520$, $\eta_p^2 = .01$, there was a small, marginally significant interaction with weapon type, $F(3, 290) = 2.32$, $p = .076$, $\eta_p^2 = .02$: across all three judgment types, those in the factual condition distinguished more clearly between weapon types than those in the original condition. In particular, participants in the factual group considered using a gun to be considerably more wrong than using a bat, $Ms = 6.32$ and 4.25, whereas in the original condition this distinction was smaller, $Ms = 5.62$ and 5.04, $F(1, 292) = 6.24$, $p = .013$, $\eta_p^2 = .02$. Neither the interaction between vignette version and outcome, $F(3, 290) = 0.64$, $p = .589$, $\eta_p^2 = .01$, nor the 3-way interaction with weapon type, $F(3, 290) = 0.15$, $p = .929$, $\eta_p^2 < .01$, approached significance.

1.2.2. Outcome, perceived negligence, and intention as predictors of moral judgments

Linear regressions indicated that, to a large extent, punishment, blame and wrongness judgments were predicted by how negligent participants considered Garmess to be (Table 2): participants who deemed Garmess to be one point more negligent judged him about three-quarters (punishment and blame) or two-thirds (wrongness) of a point more severely. The three judgments were also predicted by perceived intention, though to a lesser extent. There was no discernible outcome effect on punishment or wrongness judgments, but there was a *negative* outcome effect on blame, indicating that, when included in a model with negligence and intention, killing the family friend (i.e., the outcome was worse), was blamed slightly (.81 of a point on the 11-point scale) but significantly *less* than killing the burglar. There was no evidence of any interactions between outcome and negligence.

When the same regressions were run without the outcome x negligence interactions, no substantive differences occurred; in particular, the outcome effect remained non-significant in the punishment, $\beta = .06$, $t(294) = 1.51$, $p = .132$, and wrongness $\beta = .04$, $t(294) = 1.00$, $p = .320$, models, and unchanged in the blame model, $\beta = -.13$, $t(294) = -2.96$, $p = .003$. Similarly, including vignette version (original or factual) in the models indicated no substantive effect on punishment, $\beta < .01$, $t(294) = .02$, $p = .984$, blame $\beta = .04$, $t(292) = 1.12$, $p = .264$, or wrongness judgments, $\beta < .01$, $t(292) < .01$, $p = .998$.

1.2.3. Mediation of outcome effects on judgments by perceived negligence and intention

Mediation analyses indicated that the total effect of outcome on punishment was $c = 2.00$, $p < .001$, 95% CI [1.41, 2.59]; that is, when the victim was the friend, participants on average assigned two points more punishment than when he was a burglar. Including perceived negligence and intention in this model as possible mediators suggested that the indirect effect of outcome through perceived negligence was $a_1 * b_1 = 1.66$, $p < .001$, 95% CI [1.22, 2.14], such that the direct effect of outcome was low and non-significant, $c' = .31$, $p = .137$, 95% CI [−0.10, 0.73], and that the indirect outcome effect through intention was negligible, $a_2 * b_2 = .02$, $p = .050$, 95% CI [−0.05, 0.10] (Fig. 3). While it is possible that punishment precedes perceived negligence (see the General Discussion for consideration of causal direction), this analysis indicates that the outcome effect on punishment was almost entirely mediated by perceived negligence. If so, the reason participants assigned more punishment when the victim was a friend was because they considered Garmess to be more negligent than when his victim was a burglar.

The total effect of outcome on blame judgments was $c = 1.12$, $p = .001$, 95% CI [0.44, 1.79], and, when included in a model with perceived negligence and intention as possible mediators, its direct effect was $c' = -0.59$, $p = .028$, 95% CI [−1.12, −0.07]. This analysis indicated that the indirect effect of outcome through negligence, $a*b = 1.68$, 95% CI [1.19, 2.21], more than accounted for the total effect; that is, when negligence was held constant, the outcome effect on blame judgments was negative. There was no evidence of an indirect effect through intention, $a*b = 0.03$, 95% CI [−0.08, 0.14].

The equivalent mediation model indicated that the total effect of outcome on wrongness was $c = 1.76$, $p < .001$, 95% CI [1.17, 2.45], but that this was almost entirely accounted for by its indirect effect through negligence, $a*b = 1.55$, 95% CI [1.10, 2.06], such that its direct effect was non-significant, $c' = 0.22$, $p = .383$, 95% CI [−0.27, 0.71]. There was no indirect effect through intention, $a*b = 0.04$, 95% CI [−0.09, 0.18].

1.3. Discussion

There was a moderate outcome effect on perceived negligence: Garmess was considered more negligent when he killed the family friend rather than the burglar (the outcome IV). He was also deemed to be more negligent when he used a gun rather than a bat (the negligence IV), although this weapon effect was considerably greater when the story was presented as fact rather than as Garmess' testimony. And there was an interaction between these factors: the influence of outcome on perceived negligence was greater when Garmess used a bat rather than a gun.

The influence of outcome on the three moral judgments showed similar patterns to its influence on perceived negligence: Garmess was

Table 2
Linear Models: predictors of punishment, blame, and wrongness judgments (Experiment 1, N = 300).

|	Punishment						Blame						Wrongness					
	b	SE	β	t	p	95% CI	b	SE	β	t	p	95% CI	b	SE	β	t	p	95% CI
Constant	4.49	0.18		25.13	.000	4.14, 4.84	5.44	0.22		24.92	.000	5.00, 5.93	4.65	0.20		22.87	.000	4.23, 5.05
Outcome (burglar or friend killed)	0.29	0.21	.05	1.39	.167	−0.12, 0.75	−0.81	0.26	−.13	−3.11	.002	−1.25, −0.33	0.24	0.24	.04	1.00	.317	−0.22, 0.72
Perceived negligence	0.66	0.04	.74	17.68	.000	0.58, 0.73	0.70	0.05	.73	15.31	.000	0.59, 0.80	0.64	0.04	.68	15.02	.000	0.54, 0.73
Perceived intention	0.10	0.03	.11	2.96	.003	0.04, 0.17	0.14	0.04	.15	3.43	.001	0.05, 0.24	0.16	0.04	.16	3.96	.000	0.08, 0.23
Outcome*negligence	0.08	0.07	.04	1.11	.267	−0.07, 0.21	0.13	0.08	.06	1.59	.113	−0.02, 0.29	−0.01	0.08	.00	−0.13	.900	−0.17, 0.16
Maximum Cook's distance				.05						.10						.05		
Standardized residuals range				−2.74, 3.19						−3.06, 3.26						−3.06, 2.79		
R^2				.65						.56						.60		
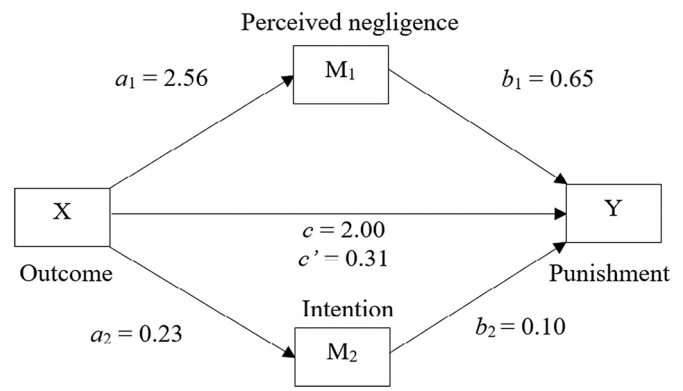
Fig. 3. Mediation model with perceived negligence and intention included as possible mediators of the outcome effect on punishment. (Experiment 1).

judged more punishable and wrong, and to a lesser extent more blameworthy, when he killed the friend than when he killed the burglar, and this outcome effect was more marked when he used a bat than when he used a gun.

The multiple regressions indicated the reason for the close similarities between the patterns of perceived negligence and punishment, blame and wrongness judgments: all three judgments were strongly predicted by perceived negligence, and none was independently predicted by outcome. This interpretation was consistent with the results of the mediation analyses, which suggested that Garmess was assigned more punishment, blame and wrongness when he killed the friend rather than the burglar because killing the friend was considered more negligent. Together, these analyses indicate that the outcome effect on moral judgments – and hence Garmess' moral luck – occurred not because of an outcome bias per se, but because he was deemed more negligent when he killed the friend than when he killed the burglar. This differential attribution of negligence might result from a hindsight bias (Kneer & Machery, 2019; Kneer & Skoczeń, 2023; Margoni, Geipel, Hadjichristidis, & Surian, 2023), but in the General Discussion we suggest an alternative explanation, according to which participants update their assessments of Garmess' actions when they are presented with the outcome.

Enzle and Hawkins (1992) reported an outcome effect on blame judgments (they did not ask about punishment or wrongness) only when Garmess used a gun, and Mazzocco et al. (2004) found an outcome effect on perceived negligence and on punishment and blame judgments (they did not ask about wrongness) whether he used a gun or a bat. We also found an outcome effect on perceived negligence and judgments, but in contrast to both Enzle and Hawkins and Mazzocco et al., this was greater when Garmess used a bat.

The reason, or reasons, for these differences between the three studies is unclear. One likely contributor to the explanation is the contrasting attitudes towards, and familiarity with, firearms in the three countries in which they were conducted; gun ownership per hundred people in the US is over 100, approximately 35 in Canada, and in the UK about 5 (Gunpolicy.org, 2020).[4] Garmess' having a gun in the house, let alone using it, is likely to have struck many of the predominantly British participants in the current study as both remarkable and remarkably dangerous. However, owning and using a baseball bat is probably also considered more serious in Britain than in Canada or the US because they are kept and used almost exclusively as weapons; almost no-one plays baseball because cricket is a much better game, and many

---

[4] In 2018 there were 13,958 firearm homicides in the US (National Center for Injury Prevention and Control, 2019). In England and Wales in 2017–18, with a population of nearly 60 million, there were 24 (Office for National Statistics, 2020).

criminals own baseball bats because guns are so difficult to acquire. These differences might therefore explain why our participants judged Garmess to be more blameworthy than did Mazzocco et al.'s US participants in all four conditions (cf. Figs. 1b and 2c), and Enzle and Hawkins' Canadians when he shot the burglar. It is also possible that Britons rated him equally blameworthy when he killed Drake the friend with a bat as with a gun because of a ceiling effect; perhaps using a gun *was* considered more blameworthy, but for some reason participants in this condition were reluctant to give more than an average rating of 7. Indeed, across all three studies, the maximum mean rating of blame was about 7 on the 0–10 scale.[5]

The two-process model (Cushman, 2008; Cushman et al., 2013) leads to predictions of a moderate outcome effect on punishment and blame, but none on wrongness. Regarding negligence, the model either makes no predictions or indicates that it would have little or no influence on judgments. In contrast, the present findings show substantial effects of perceived negligence on judgments of punishment, blame and wrongness, and that the modest outcome effects on all three occurred only because participants considered Garmess more negligent when he killed the friend rather than the burglar. However, the finding that all three judgments were influenced by participants' perceptions of Garmess' intention is consistent with this model's predictions.

According to the Path Model (Malle et al., 2014), outcome has little or no influence on blame judgments; we found that Garmess was blamed slightly more when he killed the family friend. Also according to this model, blame is assigned when agents accidentally cause harm despite having the capacity to foresee and prevent it; that is, when they are negligent. Our findings are consistent with this explanation since they indicate that perceived negligence has a substantial impact on judgments, and that the outcome effect occurred because Garmess was judged to be more negligent, and therefore more blameworthy, when he killed the friend.

These results support the negligence-based account, according to which moral judgments of accidents are influenced primarily by perceptions of negligence. As was predicted from this perspective, there was an outcome effect, but – as the regression models and mediation analyses indicate – only because participants considered killing the friend to be more negligent than killing the burglar. However, Nobes and Martin's (2022), Kneer and Machery's (2019), and Kneer and Skoczeń's (2023) finding that punishment judgments were strongly influenced by outcome as well as by negligence was not replicated.

When participants were given the original, testimonial, versions of the vignettes the weapon effect on perceived negligence and judgments was very modest. This suggests that, in the original studies and condition here, many participants doubted Garmess' testimony. They might have considered his claims that "he saw a shadow move to the left inside the doorway", or "there was no time to think things out" to be lies by which he attempted to justify his reckless behaviour. For these participants, Garmess was negligent simply because, regardless of weapon type, he could, and should, have acted differently to avoid killing his victim.

2. Experiment 2

In Experiment 1 the total effects of outcome on moral judgments were moderate or low, and its direct effects near-zero. Mazzocco et al. (2004, p. 133) reported a similarly modest outcome effect on blame judgments of $\eta^2 = .07$, compared with ours of $\eta_p^2 = .04$. One interpretation of these findings is that moral judgments are not influenced much, and perhaps not influenced directly at all, by outcomes. However, it is

also possible that this apparently small effect reflects the relatively minor contrast between the two outcomes; although killing a friend is clearly negative, many participants felt that killing a burglar is, too. If so, then to properly test the effect of outcomes on moral judgments, the distinction between outcomes needs to be much clearer. Instead of having two negative outcomes, in Experiment 2 we contrasted the very negative outcome – killing Drake the friend – with a neutral outcome in which Drake the burglar escaped and no harm was done. To avoid the possibility that some participants might consider the burglar's theft of the family's property, or his escape, to be a negative outcome, they were told that Drake dropped everything, and that he was arrested by the police a few days later. In this way the vignettes in Experiment 2 more closely resembled those of Nobes and Martin (2022) who reported much stronger outcome effects on punishment judgments when there was a stark contrast between the negative outcome - when the car ran over children - and the neutral one, when it ran over sticks. According to the models of moral judgment that emphasize the influence of outcomes on judgments – in particular the CCM and two-process model – this change should result in a much stronger direct (i.e., unmediated) outcome effect. There might also be a stronger interaction between outcome and negligence, as Nobes and Martin found with punishment judgments.η

Another possible reason for the modest or negligible outcome effects in Experiment 1 was that this was a discrete, binary variable; Garmess killed either a friend or a burglar. In contrast, there were two negligence variables: the binary IV – he used either a gun or a bat – and the continuous perceived negligence covariate. Presumably perceptions of negligence are more direct influences on judgments than the IV, and this variable is likely to have conferred more statistical power than the binary IV because it was continuous. It is possible that at least some of the disparity in effect sizes between perceived negligence and outcome arose for these reasons. To address this issue, in Experiment 2 we asked participants to rate the severity of the outcome on an 11-point scale. Another advantage of including this question was that it provided an extra comprehension check: a participant who said that killing a friend was not a serious outcome, or that causing no harm was, is likely to have misunderstood the vignette or question.

Just as the outcome was poorly manipulated in the original studies and in Experiment 1, so too was negligence. In Enzle and Hawkins' (1992) and Mazzocco et al.'s (2004) studies negligence was manipulated by having Garmess use either a bat or a gun, but we found that, especially in the original (testimonial) condition, taking these to investigate noises was considered to be almost equally negligent. In fact, the (relatively slight) contrast between outcomes influenced perceived negligence considerably more than did the weapon effect. However, we did find a strong effect of *perceived* negligence – as opposed to the manipulated negligence IV – on all three judgment types. This suggests that, regardless of whether Garmess used a gun or a bat, his actions could be considered highly negligent; for example, he did not try to identify Drake by calling out, and he made no attempt to find out if Drake was armed. Many participants are likely to have thought that Garmess' actions were reckless and dangerous, which led him to commit a homicide. To increase the weapon effect, in Experiment 2 we increased the contrast between the negligence IV levels by introducing a third level in which Garmess was unarmed. We adapted one of Mazzocco et al.'s (2004) vignettes so that, in this low negligence condition, Garmess merely called out to the suspected intruder.[6] Drake was startled and, losing his footing, fell down the stairs. When he was a family friend, he hit his head and died, but when he was a burglar, he was unharmed.

For two reasons we used the factual versions of the vignettes, rather than the original, testimonial, versions. First, we found in Experiment 1 a much clearer effect of weapon type when the factual vignettes were

---

[5] In Experiment 2 the pattern of blame ratings was similar to that reported by Mazzocco et al. in that there was an outcome effect when Garmess used a gun as well as when he used a bat. This is consistent with there being a ceiling effect in Experiment 1 which was attenuated somewhat in Experiment 2 by the sole use of the factual version of the vignettes.

[6] Mazzocco et al. (2004, Study 4) did not manipulate weapon type in this study and so did not compare judgments when Garmess was armed and unarmed.

presented, and this version worked at least as well as the testimonial version on all other counts. Second, the factual versions worked better with the new neutral outcome because it would be strange for Garmess to give testimony – as if he were in court – when no harm was done to Drake.

The two-process model in particular stresses the influence of causal responsibility on judgments of punishment and blame. It is possible that the substantial effects of perceived negligence in Experiment 1 actually reflected participants' perceptions of causality. In Experiment 2 we therefore asked participants to rate the extent to which they considered Garmess to have caused the outcome. The two-process model would predict that causal responsibility is the primary influence on punishment and blame judgments of accidental agents, whereas the negligence-based account predicts that negligence remains the main influence even when causal responsibility is included in the analyses.

2.1. Method[7]

2.1.1. Design

The two main differences from Experiment 1 were that there were three negligence IV levels – gun and bat (as in Experiment 1), and no weapon – and the two outcome IV levels were the innocent friend being killed (as in Experiment 1), and the burglar not being harmed (instead of being killed). In this experiment only the factual versions of the vignettes were used. In addition, participants were asked to rate the severity of the outcome and the extent to which Garmess caused it. There was also an open question asking them to justify their first judgment. This provided a comprehension and engagement check.

2.1.2. Recruitment and sample

A total of 185 people responded to invitations sent to university participant panels, and 118 via Amazon's Mechanical Turk (MTurk). Of these 303, 23 (7.6%) did not start the questionnaire or provided no relevant data, 10 (3.0%) opted out, and 29 (9.6%; 24 from MTurk) were excluded because their responses indicated a lack of comprehension or engagement (for example, saying that killing an innocent friend was not at all serious, or that calling for Drake to stop, who then ran away unharmed, should be punished with more than ten years' imprisonment).

Sensitivity analyses (Perugini et al., 2018) using G*Power (Faul et al., 2007) for the MANOVA (6 groups, 3 DVs) and linear multiple regressions (5 predictors) indicated that the sample of 241 participants was sufficient to detect even small effect sizes (f^2s = .044 and .084 [η_p^2s = .042 and .077], respectively) with α = .05 and power = .95.

The participants were all 18 or over, their mean age was 32.1 years, and 161 (66.7%) were women. 217 (90.0%) had continued education at least until 18. 136 (56.4%) described themselves as British, 76 (31.5%) American, and 11 (4.6%) European. English was the first language of 220 (91.3%).

2.1.3. Materials and procedure

The same factual versions of the vignettes were used as in Experiment 1, except that an additional one was adapted from Mazzocco et al. (2004, Study 4) in which Garmess was unarmed and merely called for Drake to stop. This startled Drake, who lost his footing and fell down the stairs. The other change was that, instead of Drake the burglar being killed, the outcome was neutral because he now ran off, unharmed. The negative outcome remained that Drake the family friend was killed by the gun or bat (as in Experiment 1), or, when Garmess was unarmed, by hitting his head on a stair as he fell.

As well as the judgments, negligence and intention questions that were asked in Experiment 1, after their first judgment participants were asked the open question, "Why did you make the response that you did?"

to check their comprehension. They were also asked about causation, "To what extent did Garmess cause things to turn out the way they did?", and about the severity of the outcome, "How severe (serious) was the outcome?" As with all the questions in Experiment 1, ratings were made on 11-point scales, from "Garmess didn't cause the outcome at all" to "It was entirely caused by Garmess", and from "Not at all serious or severe" to "Extremely serious or severe".

2.2. Results[8]

2.2.1. The influence of the manipulated IVs (negligence and outcome) on perceived negligence and judgment ratings

A 3 (weapon type: [none, bat, gun]) x 2 (outcome [no harm, friend's death]) between subjects ANOVA indicated a strong main effect of weapon type (the negligence IV) on perceived negligence, $F(2, 229) = 97.58$, $p < .001$, $\eta_p^2 = .46$: Garmess' negligence was considered very low when he used no weapon, $M = 1.31$ on the 0–10 negligence scale, more when he used a bat, $M = 5.06$, and most when he used a gun, $M = 6.96$ (Fig. 4a). Bonferroni post hoc tests indicated that all three comparisons were significant, $ps < .001$. There was also a lesser effect of outcome, $F(1, 229) = 41.54$, $p < .001$, $\eta_p^2 = .15$; participants considered Garmess more negligent when he killed Drake the friend than when he disturbed Drake the burglar, $Ms = 5.45$ and 3.41. However, these main effects were qualified by an interaction between weapon type and outcome, $F(2, 229) = 14.77$, $p < .001$, $\eta_p^2 = .11$; killing the friend was attributed more negligence than disturbing the burglar when Garmess used a gun, $F(1, 76) = 23.34$, $p < .001$, $\eta_p^2 = .24$, or a bat, F(1, 80) = 32.33, $p < .001$, $\eta_p^2 = .29$, but not when he did not have a weapon, $F(1, 79) = .88$, $p = .351$, $\eta_p^2 = .01$.

An equivalent MANOVA with all three judgments as the DVs indicated a strong main effect of outcome, $F(3, 226) = 54.93$, $p < .001$, $\eta_p^2 = .42$, a main effect of weapon type, $F(6, 454) = 26.22$, $p < .001$, $\eta_p^2 = .26$, and an interaction between the two, $F(6, 454) = 23.87$, $p < .001$, $\eta_p^2 = .13$.

Outcome influenced all three judgments, especially punishment (Figs. 4b-d): When Garmess killed the friend, he was judged considerably more punishable, $F(1, 228) = 156.00$, $p < .001$, $\eta_p^2 = 0.41$, more wrong, $F(1, 228)$, $p < .001$, $\eta_p^2 = .26$, and more blameworthy, $F(1, 228) = 41.55$, $p < .001$, $\eta_p^2 = .15$, than when he disturbed the burglar. Weapon type also strongly influenced all three judgments: punishment, $F(2, 228) = 66.09$, $p < .001$, $\eta_p^2 = .37$; blame, $F(2, 228) = 70.81$, $p < .001$, $\eta_p^2 = .38$; and wrongness, $F(2, 228) = 92.27$, $p < .001$, $\eta_p^2 = .45$.

These main effects were all qualified by weapon x outcome interactions. Regarding punishment, $F(2, 228) = 35.03$, $p < .001$, $\eta_p^2 = .24$, there was a strong outcome effect when Garmess used a gun, $F(1, 76) = 98.35$, $p < .001$, $\eta_p^2 = .56$, and a bat, $F(1, 80) = 92.17$, $p < .001$, $\eta_p^2 = .54$, but not when he had no weapon, $F(1, 79) = 1.21$, $p = .275$, $\eta_p^2 = .02$.

The interaction was less substantial with wrongness judgments, $F(2, 228) = 15.44$, $p < .001$, $\eta_p^2 = 0.12$. There was an outcome effect when Garmess used a gun, $F(1, 75) = 32.47$, $p < .001$, $\eta_p^2 = .30$, and a bat, $F(1, 80) = 51.07$, $p < .001$, $\eta_p^2 = .39$, but when he had no weapon, outcome made little or no difference, $F(1, 79) = 3.34$, $p = .071$, $\eta_p^2 = .04$.

Regarding blame judgments, the weapon x outcome interaction was smaller still, $F(2, 228) = 5.95$, $p = .003$, $\eta_p^2 = .05$. Killing a family friend was considered more blameworthy than disturbing a burglar when the weapon was a gun, $F(1, 76) = 34.02$, $p < .001$, $\eta_p^2 = .31$, and when it was a bat, $F(1, 80) = 18.03$, $p < .001$, $\eta_p^2 = .18$, but not when Garmess did not use a weapon, $F(1, 79) = 2.14$, $p = .147$, $\eta_p^2 = .03$.

2.2.2. Perceived severity, negligence, causal responsibility and intention as predictors of moral judgments

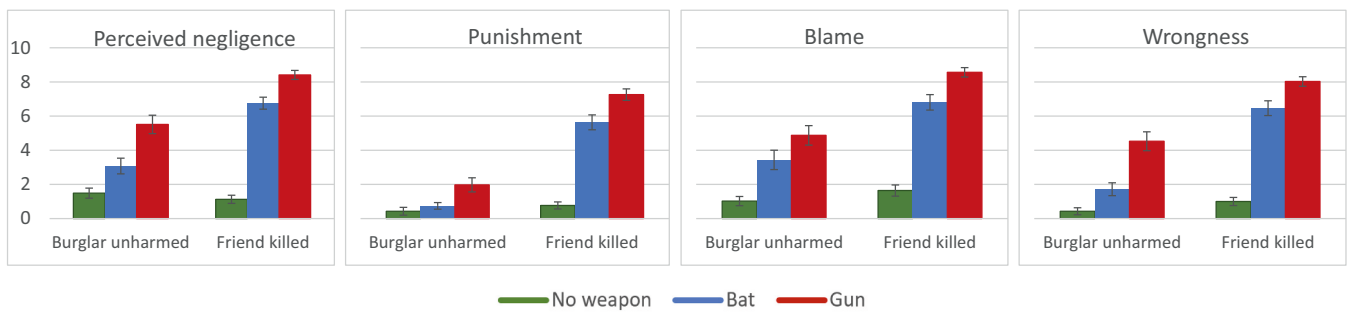Linear models of predictors of the three judgments are shown in

---

Fig. 4. Mean [+ SEs] ratings of a) perceived negligence, b) punishment, c) blame, and d) wrongness (0 = not at all / none; 10 = very high) by outcome (burglar unharmed or friend killed) and weapon type (none, bat or gun). (Experiment 2.)

Table 3. Punishment judgments were predicted by both the perceived severity of the outcome, and by the extent to which participants considered Garmess to be negligent. There was also an interaction between these factors: high levels of punishment were typically assigned when, and only when, Garmess was considered negligent *and* he killed the innocent family friend. The extents to which he was judged causally responsible, and to which he was thought to have intended the outcome, were also weaker predictors of punishment.

In contrast, both blame and wrongness judgments were largely predicted by negligence, and the outcome effect, though significant, was considerably smaller. The interaction between negligence and outcome was also weak, and in the case of blame, not significant. As with punishment, causal responsibility and intention were also modest predictors of blame and wrongness.

2.2.3. Mediation of outcome effects on judgments by perceived negligence, causal responsibility and intention

Mediation analyses indicated that the total effect of outcome severity on punishment was $c = 0.44$, $p < .001$, 95% CI [0.35, 0.53], that is, participants who considered the outcome to be one point more severe judged Garmess to deserve nearly half a point more punishment. Including perceived negligence, causal responsibility and intention in this model as possible mediators suggested that the direct effect of punishment was $c' = 0.24$, $p < .001$, 95% CI [0.17, 0.31], its indirect effect of outcome on punishment through perceived negligence was $a_1*b_1 = 0.12$, 95% CI [0.07, 0.18] and through causal responsibility was $a_2*b_2 = 0.06$, 95% CI [0.03, 0.11]. The indirect effect through intention accounted for little or none of the outcome effect, $a_3*b_3 = 0.02$, 95% CI [0.00, 0.04]. While it is possible that punishment precedes perceived negligence and causal responsibility (see the General Discussion for consideration of causal direction), this model indicates that, to some extent, participants assigned more punishment to Garmess when he killed the friend because of this outcome per se, but also because they considered him more negligent and causally responsible.

The total effect of outcome severity on blame judgments was $c = 0.35$, $p < .001$, 95% CI [0.24, 0.46]. Including perceived negligence and intention in the model as possible mediators indicated that the indirect effect of outcome through negligence was $a_1*b_1 = 0.17$, 95% CI [0.10, 0.24], through causal responsibility, $a_2*b_2 = 0.10$, 95% CI [0.05, 0.15], and through intention, $a_3*b_3 = 0.02$, 95% CI [0.00, 0.05]. The direct effect of outcome was $c' = 0.07$, $p = .077$, 95% CI [−0.01, 0.14]. This suggests that participants judged Garmess to be more blameworthy when he killed the friend largely or solely because they considered him more negligent and causally responsible.

The equivalent mediation model indicated that the total effect of outcome severity on wrongness judgments was $c = 0.37$, $p < .001$. 95% CI [0.27, 0.48], and that its indirect effect was primarily through negligence, $a_1*b_1 = 0.21$, 95% CI [0.13, 0.29] and to some extent causal responsibility, $a_2*b_2 = 0.06$, 95% CI [0.03, 0.10], but only slightly or not at all through intention, $a_3*b_3 = 0.01$, 95% CI [0.00, 0.03]. The direct effect of outcome on wrongness was $c' = 0.10$, $p = .001$, 95% CI [0.04,

0.15]. This model suggests that the outcome effect on wrongness was largely explained by participants considering Garmess more negligent and – less so – causally responsible, but that there was also a modest unmediated influence of perceived outcome severity.

2.3. Discussion

Two main changes were made in this experiment: first, to increase any outcome effects, the contrast between outcomes was raised so that Garmess either killed a friend or harmlessly disturbed a burglar. And second, to increase any negligence effects, the contrast between negligence levels was raised by adding another level so that Garmess carried a gun or a bat, or was unarmed.

The change to the neutral outcome so that Garmess now harmlessly disturbed rather than killed a burglar had little or no effect on ratings of perceived negligence, blame and wrongness; these were similar across the two experiments (Figs. 2a and 4a). However, in Experiment 2 there was less evidence of interactions between weapon type (bat or gun) and outcome: use of the gun was considered more negligent, blameworthy and wrong than use of the bat, regardless of the outcome.

In contrast, the pattern of punishment judgments in Experiment 2 was influenced by the change to a neutral outcome: when Garmess merely disturbed the burglar, the large majority of participants assigned him little or no punishment almost regardless of how negligent they considered him to be.

The added unarmed level of negligence strongly influenced perceived negligence and all three judgments. In particular, all outcome effects disappeared: when he did not use a weapon, Garmess was assigned low levels of negligence, punishment, blame and wrongness, regardless of the outcome.

Consistent with the two-process model (Cushman, 2008; Cushman et al., 2013; Martin & Cushman, 2016), the extent to which Garmess was considered causally responsible for the outcome influenced punishment and blame judgments. This factor also partially accounted for the outcome effects on these judgments. However, its effect sizes were considerably lower than those of negligence, which remained the primary influence on judgments even when causal responsibility was included in the models. In addition, according to the two-process model, causal responsibility influences punishment and blame but not wrongness judgments, but it had approximately the same effect on all three.

3. General discussion

Two experiments were conducted to investigate the relative influence of outcome and negligence on moral judgments of accidental agents. Strong outcome effects (i.e., moral luck) have been widely reported in previous research (e.g., Bornstein, 1998; Gino et al., 2010; Kurdi et al., 2020; Martin & Cushman, 2015; Mazzocco et al., 2004; McNamara et al., 2019; Patil et al., 2017; Robbennolt, 2000; Young et al., 2010), but negligence has received considerably less attention. We aimed also to examine whether the outcome effects might be partially or

Table 3
Linear Models: predictors of punishment, blame, and wrongness judgments (Experiment 2; N = 241).

	Punishment						Blame						Wrongness					
	b	SE	β	t	p	95% CI	b	SE	β	t	p	95% CI	b	SE	β	t	p	95% CI
Constant	1.32	0.28		4.73	.000	0.71, 1.89	2.41	0.37		6.51	.000	1.54, 3.27	2.46	0.29		8.58	.000	1.79, 3.18
Outcome severity	0.28	0.03	.35	10.00	.000	0.22, 0.34	0.09	0.04	.10	2.33	.021	0.00, 0.18	0.11	0.03	.12	3.57	.000	0.04, 0.17
Perceived negligence	0.44	0.04	.48	10.91	.000	0.36, 0.53	0.60	0.05	.57	11.19	.000	0.48, 0.72	0.73	0.04	.71	17.48	.000	0.62, 0.84
Causal responsibility	0.15	0.04	.15	3.61	.000	0.07, 0.24	0.25	0.06	.21	4.33	.000	0.13, 0.37	0.16	0.04	.14	3.57	.000	0.07, 0.25
Intention	0.07	0.04	.07	1.85	.066	−0.01, 0.15	0.14	0.05	.13	2.81	.005	0.03, 0.26	0.08	0.04	.07	1.95	.052	−0.01, 0.17
Outcome*Negligence	0.07	0.01	.29	9.06	.000	0.05, 0.08	0.01	0.01	.05	1.37	.171	−0.01, 0.03	0.02	0.01	.07	2.22	.028	0.00, 0.03
Maximum Cook's distance				.05						.11						.08		
Standardized residuals range				−2.80, 3.07						−2.76, 3.16						−3.01, 3.07		
R^2				.78						.70						.81		

wholly explained by the differential attribution of negligence, that is, agents who unluckily cause harm being considered more negligent, and for this reason more culpable, than those who luckily cause no harm. Predictions from five models or accounts of moral judgment were tested.

Experiment 1 replicated studies conducted by Enzle and Hawkins (1992) and Mazzocco et al. (2004) in which an agent – Garmess – accidentally killed an apparent intruder – Drake. Outcome was manipulated by Drake being either a burglar or a family friend, and negligence by Garmess using either a bat or a gun. Consistent with this previous research, outcome predicted moral judgments: Garmess was assigned more punishment, blame and wrongness when his victim was a family friend rather than a burglar. However, perceived negligence also strongly predicted judgments, and regression and mediation models indicated that the outcome effect on all three judgment types occurred largely or solely because participants rated killing a friend to be more negligent than killing a burglar.

In Experiment 2 the distinction between outcomes was increased by having Garmess harmlessly disturb the burglar rather than kill him, and between levels of negligence by adding a condition in which Garmess was unarmed. As in Experiment 1, the strongest predictor of all three judgments was the degree to which Garmess was considered negligent. Causal responsibility and intention also appeared to have moderate and slight impacts, respectively. But in contrast to the first experiment, there was also a strong and direct (i.e., unmediated) outcome effect on punishment judgments, and an interaction between negligence and outcome: most participants assigned relatively severe punishment when, and only when, they considered Garmess negligent *and* the outcome was negative. Garmess was also considered more blameworthy and wrong when he killed the friend than when he disturbed the burglar; although this was partly explained by an indirect effect through negligence, there was also a modest but significant direct effect of outcome on both these judgments.

These findings indicate that the outcome effect differs according to the type of moral judgment. All three are strongly influenced by perceived negligence, which also largely explains the outcome effect on blame and wrongness judgments: unlucky accidental agents are considered more negligent than lucky ones. But punishment judgments are different because – at least when there is a clear distinction between negative and neutral consequences, as in Experiment 2 – they are also strongly and directly influenced by outcomes. This difference between punishment judgments on the one hand, and blame and wrongness on the other, occurs primarily when accidental agents luckily cause no harm; although these agents are considered negligent, blameworthy and wrong, they are seldom assigned any punishment. In contrast, when agents unluckily cause harm, the pattern of punishment judgments closely follows that of negligence, blame and wrongness. These findings therefore indicate that the large majority of people assign punishment according to a conjunction rule: punish if, and only if, the accidental agent is negligent *and* the outcome is negative. The implication is that moral luck applies more to punishment than to blame and wrongness judgments, and specifically to lucky rather than to unlucky agents. Nobes and Martin (2022) report a similar imbalance in moral luck.

The main difference between the findings of the two experiments was that, whereas in Experiment 1 punishment judgments closely resembled perceived negligence and blame and wrongness judgments, in Experiment 2 they did not. This was because Garmess was considered punishable in Experiment 1 when he killed the burglar, but he was assigned little or no punishment in Experiment 2 when the burglar was unharmed. It is explained by both outcomes in Experiment 1 – killing a friend and killing a burglar – being considered negative, resulting in the assignment of punishment. In contrast, when the burglar was unharmed in Experiment 2, the outcome was neutral and so there was no conjunction of negative outcome and negligence, and, as a result, punishment was not assigned.

3.1. Negligence-based account

Regarding blame and wrongness judgments of accidental agents, the present findings are consistent with the predictions of the negligence-based account that perceptions of negligence are the principal influence (proposal 1), and that outcome influences these judgments because it influences people's assessment of negligence (proposal 3). However, the negligence-based account did not predict the strong and direct outcome effect that we found on punishment judgments in Experiment 2. This experiment's findings therefore closely replicate those of Nobes and Martin (2022) that also had starkly contrasting negative and neutral outcomes; children were either killed or unharmed.

These results also corroborate Kneer and Machery's (2019) and Kneer and Skoczeń's (2023) central findings that negligence strongly influences punishment, blame and wrongness judgments, and that, while there is also a strong direct outcome effect on punishment, any influence of outcome on blame and wrongness judgments occurs largely or solely through negligence. Moreover, by manipulating the negligence IV (weapon type) as well as the outcome, we were able to extend this analysis and show the interaction between outcome and negligence regarding punishment judgments. This manipulation also enabled us to test and empirically support Kneer and colleagues' assumption that negligence is causally antecedent to moral judgment, rather than being a post hoc rationalization of outcome-based judgment.

3.2. Path Model of blame

According to the Path Model (Malle et al., 2014), an agent should be blamed when they unintentionally cause harm that they have an obligation and the physical and cognitive capacity to prevent. Here, Garmess was obliged in particular not to harm the friend, and he had the capacity to avoid doing so. As discussed in the Introduction, Malle et al. (p. 155) explain that this cognitive capacity is closely related to the concept of negligence, and so their model is largely supported by the present findings of negligence strongly influencing blame judgments. Malle and colleagues also play down the role of outcomes in blame judgments, and this point too was at least partially supported: while there were outcome effects in both experiments, they were largely or solely accounted for by differences in perceived negligence and causal responsibility. The similarities between blame and wrongness judgments reported here suggest that a similar analysis to the Path Model can be extended to wrongness, too. However, it is probably fair to say that these researchers' view that outcome effects on blame judgments are minimal (Malle et al., 2014, pp. 160–162) extends also to punishment judgments, and that therefore they would not have predicted the strong outcome effect on punishment judgments that we found in Experiment 2.

3.3. Enzle and Hawkins

The key prediction from Enzle and Hawkins (1992) was that there would be an interaction between outcome and negligence such that high levels of blame would be assigned only when Garmess was negligent and the outcome was negative. When we replicated this study in Experiment 1, our finding of a stronger outcome effect when Garmess used a bat contrasted with theirs that the outcome effect occurred only when Garmess used a gun (Figs. 1a and 2c). However, in Experiment 2 we added a level of non-negligence (Garmess was unarmed) and of no harm (the burglar escaped) and, as Figs. 1a and 4c illustrate, our results resemble theirs: there was an outcome effect on blame judgments only when Garmess was negligent, that is, when he was armed. We found this also to be the case with punishment and wrongness judgments (Figs. 4b and d). We therefore agree with Enzle and Hawkins' central claim that, "When people are not negligent, and are hence blameless at the initiation of an act, they will be relatively blameless for the consequences of the act. Blame *will* attach to consequences, however, when the initiation of the act involves substantial negligence" (p. 172). However, our agreement is limited to when distinctions between outcomes and between levels of negligence are very clear (i.e., negative vs. neutral), as in Experiment 2. Also, whereas punishment is strongly "attached" to consequences – even when agents are negligent it is almost invariably assigned only when the outcome is negative – blame and wrongness judgments are less influenced by outcomes, and primarily only because actions with worse outcomes are considered more negligent.

3.4. Culpable Control Model and the direction of causality

The main prediction from Alicke and colleagues' CCM (e.g., Alicke, 2000; Alicke, 2014) is that there is "No constraint by prior culpability" (Mazzocco et al., 2004, p. 131), that is, the outcome bias occurs independently of negligence. Comparison of Figs. 1b and 4c indicates similarities between their findings and ours. In particular, whether Garmess used a bat or a gun the outcome effect was very similar: he was considered more blameworthy (and punishable and wrong) when he killed the friend. However, for two reasons, their prediction was not supported. First, we found no outcome effect when Garmess was unarmed and therefore not negligent. This is particularly striking given that the contrast between outcomes was much starker, and therefore the outcome effect should have been much stronger, than in Mazzocco et al. (and in Enzle and Hawkins, and Experiment 1 here). And second, our analyses indicate that the principal reason for outcome effects on moral judgments was that actions with more negative outcomes were deemed to be more negligent. The only exception was that punishment judgments in Experiment 2 were also directly influenced by outcome, but this was moderated (i.e., constrained) by negligence; participants assigned Garmess more punishment when he killed the friend only when they considered him negligent.

According to Alicke and colleagues and other proponents of motivated-blame models of moral judgment (e.g., Alicke, 2000, 2014; Greene, 2008; Haidt, 2001; Mazzocco et al., 2004), people want to blame harm-doers, and so their judgments are intuitive, spontaneous, and outcome-based. Negligence ratings are merely post hoc rationalizations, and so perceived negligence is influenced by judgments. This direction of causality (outcome → judgments → perceived negligence) contrasts with that suggested by the negligence-based account, according to which blame and wrongness judgments of accidental harms are primarily negligence-based, and influenced only slightly and indirectly (through negligence) by outcome (outcome → perceived negligence → judgment). Regarding punishment, the current findings support the motivated-blame approach insofar as these judgments are directly influenced by outcome, but they also indicate that this outcome effect is moderated by negligence; that is, punishment judgments are based primarily on negligence *and* outcome (outcome & perceived negligence → judgment).

As a result of the associations between the perceived negligence and judgments variables being only correlational, the regression and mediation analyses alone cannot help to establish whether the current findings support the motivated-blame prediction that judgments influence perceived negligence, or the negligence-based prediction that perceived negligence influences judgments. However, because in the current studies the negligence IV was manipulated experimentally, we can be confident that it had a clear and independent effect on perceived negligence. This was shown by the ANOVAs (which included only the manipulated IVs as predictors), and was particularly clear in Experiment 2, when Garmess' negligence was much lower when he was unarmed than when he took a bat or gun. The independence of negligence effects on judgments is also demonstrated by different levels of negligence being judged differently even within levels of outcome, and by punishment judgments being influenced separately and independently by *both* outcome and negligence. These findings are consistent with those of Nobes and Martin (2022), who used the same perceived negligence question as the one used here, but manipulated negligence in a very

different way. In addition, all these findings resonate with cases such as the hypothetical drivers described in the Introduction who had identical intentions and accidentally caused identical outcomes, but who were judged very differently because of their different levels of negligence. Together, all of these points provide strong support for the negligence-based account's proposal that negligence influences judgments, rather than the motivated-blame claim that judgments influence perceived negligence.[9]

However, as discussed below, these two approaches are not necessarily mutually exclusive, and it is possible that influences between judgments and negligence are bidirectional. While the more parsimonious explanation of the current findings is that negligence affects judgments, the possibility that judgments also exert some effect on perceived negligence cannot be ruled out.

3.5. Two-process model

Like the CCM, the two-process model (Cushman, 2008; Cushman et al., 2013) leads to predictions of strong outcome effects on punishment and blame judgments that are independent of negligence. These received little or no support in Experiment 1 because, although killing the friend was assigned more punishment and blame than killing the burglar, these outcome effects occurred largely or solely because killing the friend was considered more negligent. However, in Experiment 2 the prediction regarding punishment was partially supported because there was a strong and direct outcome effect, although only when Garmess was negligent. Regarding blame judgments, support for the two-process model was very limited: the influence of negligence was much stronger than that of outcome, and even this small outcome effect was largely accounted for (i.e., mediated) by negligence.

The two-process model also predicts a strong influence of causal responsibility on punishment and blame. There was some support for this in that causal responsibility had an effect on all three judgments, although this was considerably less than the influence of negligence.

With regard to wrongness judgments, the two-process model predicts little or no influence of negligence or outcome. Our findings indicate that in both experiments the impact of negligence was particularly strong. The two-process model also predicts a strong effect of intention on wrongness, but in both experiments its influence was only modest. However, it is important to qualify this last point because we focused on accidental actions, and intention effects were low because almost all participants in all conditions considered Garmess' intentionality to be low or nil.

Cushman (2008) reported that punishment and blame judgments were similar, in that both were influenced by intention and outcome, and that wrongness was different because it was influenced only by intention. In contrast, Kneer and Machery (2019), Kneer and Skoczeń (2023), and Nobes and Martin (2022) found that blame and wrongness judgments were similar, and that punishment alone was strongly and directly influenced by outcome. The findings of Experiment 2 reported here are consistent with the latter; blame is similar to wrongness rather than to punishment.

The reasons for these discrepant findings concerning blame might arise from the participants interpreting the questions differently. Whereas Kneer and colleagues and we asked explicitly how blameworthy the agent was for their *action*, Cushman's (2008) question – "How much blame does [agent] deserve?" – might have been interpreted in terms of blame for the *outcome*. Whereas responses to the former are, like wrongness judgments, unaffected by the outcome, responses to the latter are, like punishment judgments, dependent on it. For example, when judging a reckless driver who luckily causes no harm, we would

say that they are wrong and blameworthy for driving recklessly (their action), but blameless and not punishable for the harm they caused (the outcome) because there *was* no harm. (Similarly, we would say that the driver and driving were irresponsible, but that the driver was not responsible for any harm.) This would explain why, when asked explicitly about actions (as in the more recent studies), blame and wrongness judgments are similar, and when asked implicitly or explicitly about outcomes (perhaps as in Cushman, 2008), blame and punishment judgments are similar. (For an investigation of contrasting responses to agent- and outcome-focused questions, see Nobes, Panagiotaki, & Bartholomew, 2016).

3.6. The differential attribution of negligence: motivated blame, hindsight bias, and "rational" updating

The finding that, especially regarding blame and wrongness judgments, outcome effects are explained by the differential attribution of negligence, raises an important question: why should outcome influence perceived negligence? For example, why should Garmess be considered more negligent when he kills Drake the friend than when Drake the burglar escapes unharmed? Mazzocco et al. (2004) argue that basing judgments on perceived negligence, which is itself influenced by outcome, is no more rational than basing judgments directly on outcome. As discussed above, they take a motivated-blame approach and argue that judgments must be outcome-based, and that negligence ratings are merely post hoc justifications of these judgments.

In contrast, Kneer and Machery (2019) and Kneer and Skoczeń (2023) propose that the differential attribution of negligence reflects a hindsight bias (e.g., Bernstein et al., 2011; Hawkins & Hastie, 1990), that is, the tendency to consider an event to be more probable after than before it occurs (see also Margoni et al., 2023). In experiments such as theirs and ours, participants for whom the outcome is neutral will therefore judge a negative outcome to have been unlikely, and so the action to be low-risk and the agent non-negligent. Conversely, participants for whom the outcome is negative will judge this outcome to have been more probable, and so the action more risky and the agent more negligent.

Consistent with this view that outcome effects on moral judgments arise from hindsight bias, Kneer and colleagues found that participants' assessments of the probability of outcomes mediated the influence of outcome on negligence, which in turn influenced judgments. Moreover, when these researchers presented participants with both negative and neutral outcomes in their within-subjects and counterfactual priming studies (see also Lench et al., 2015), and when they gave expert opinions on risk, participants showed substantially decreased outcome effects on ratings of probability, negligence and blame, and, to a lesser extent, punishment. Kneer and colleagues interpret the latter findings in terms of these methods attenuating the hindsight bias, and therefore as demonstrating that this bias was the source of the original outcome effect.

We suggest an alternative explanation of the outcome effect on negligence assessments of accidental actions. First, consider the experience of participants in these experiments for whom Drake turned out to be a burglar. When they read that Garmess heard noises that couldn't be made by a family member, most are likely to have assumed, like Garmess, that there was a burglar upstairs. Moreover, on the basis of this assumption, many are likely to have considered it reasonable for Garmess to take a weapon in case he had to defend himself. When these participants learned the outcome – that Drake was, indeed, a burglar – Garmess' and their assumption was confirmed, and so Garmess' actions, though perhaps not condoned, were understandable and therefore not particularly negligent or culpable.

Now compare this with the experience of participants in the other condition: since the vignettes began in the same way, most of them, too, will initially have assumed that there was a burglar upstairs, and so many will have thought it reasonable for Garmess to take a weapon.

---

[9] In the Supplementary Material we report reverse causal analyses that compare the two possible pathways: outcome → judgment → negligence; and outcome → negligence → judgment. These too show support for the latter.

However, when they read that Drake was an innocent friend, these participants were given new information that strongly contradicted their assumptions, and so is likely to have led them to reassess, or update, their views and judgments of Garmess' actions. Many of these participants will now have thought that actually he took an unnecessary and unreasonable risk in arming himself without, for example, finding out who was upstairs, or shouting a warning. Taking and using a lethal weapon was therefore negligent, and so he was blameworthy and should be punished accordingly.

The key point is that the different outcomes in the two conditions provided participants with very different information. Notice that for participants in the first group there was *never any suggestion* that Drake might not be a burglar, and so the possibility of his being an innocent friend probably didn't even *occur* to them. In contrast, this possibility *must* have occurred to participants in the second group because they learned that he actually *was* an innocent friend. Our suggestion is therefore that the outcome effect on assessments of negligence – and hence on judgments of culpability – results from people basing these assessments on the information that is available to them; it follows that, when the outcomes provide very different information, people make very different judgments.

This account of outcome influencing perceptions of negligence involves similar updating to that proposed and reported by Monroe and Malle (2019). While they found that new causal and mental state information was systematically processed to influence judgments, we propose that, in the same Bayesian-like, abductive way, participants update their assessments of negligence, and hence of culpability, as outcome information is processed.

The updating account of the outcome effect on negligence is consistent with, and offers an alternative explanation of, Kneer and Machery's (2019) and Kneer and Skoczeń's (2023) finding that outcome effects on perceived negligence are mediated by ratings of the probability of outcomes. Initially, most participants shared the prevailing assumption that Drake was probably a burglar and that it was therefore reasonable and non-negligent for Garmess to arm himself; participants for whom Drake turned out to be a burglar had this assumption and assessment of probability confirmed and so continued to consider Garmess to be non-negligent; but those for whom Drake turned out to be an innocent friend had this assumption contradicted, and so they reassessed the probability of Drake's being a burglar downwards, and the probability that he was innocent upwards, and updated their ratings of Garmess' negligence accordingly.

This explanation is also in line with Kneer and colleagues' findings from within-subjects and counterfactual priming studies, that giving participants both possible outcomes attenuates the outcome effect. For example, letting participants know that Drake *could* have been an innocent friend should challenge participants' assumptions that he was a burglar in the same way as telling them that he *was* an innocent friend. Now the possibility that Drake was a friend *must* occur to all participants because they are explicitly given this information, with the result that they should judge him equally negligent regardless of the actual outcome, that is, there should be little or no outcome effect on negligence and hence on blame and wrongness.

Kneer and Machery (2019) distinguish their "hindsight bias" view from the "rationalist" approach, in which outcome information is taken as evidence for the probability of an event. In this sense, the updating explanation of the outcome effect proposed here is a rationalist approach because it is reasonable and appropriate for people to base their assessments of probability, and hence of negligence and culpability, on the best available information, and to update their assessments and judgments when the outcome changes this information. This is not hindsight *bias*, but reasoned re-evaluation *with the benefit of hindsight.*

Malle and colleagues tend to avoid the terms "rational" and "irrational" because the information processing on which blame judgments are based can be fallible and influenced by emotion. Instead, they describe the processing as "systematic" (2014, p. 152). We acknowledge

and agree with this point, but suggest that they are used in a relative, rather than absolute sense because it is more rational to base moral judgments on controllable factors such as intention and negligence, and on systematic updating of assessments of risk as new information (including from outcomes) becomes available, than on uncontrollable outcomes and cognitive distortions such as hindsight and motivated biases. These terms therefore highlight important and, we think, conceptually helpful distinctions between the various accounts of moral judgments. Similarly, Haidt (2001) contrasts rationalist, reasoning-based approaches with his social intuitionist model of moral judgment, Gino, Moore, and Bazerman (2008) distinguish between rational and intuitive mindsets, and Kim, Park, and Young (2020) compare rational with motivated impression updating.

3.7. Punishment judgments

A further question raised by these findings is, why might the direct effect of outcome be much stronger on punishment than on blame and wrongness judgments? We suggest that there are adaptive reasons for punishment focusing more on outcomes than do blame and wrongness (Martin & Cushman, 2016). Specifically, by punishing negative outcomes, even when accidentally caused, the punisher sends a signal to the transgressor and to observers that such outcomes must not be allowed to recur. Importantly, it is unlikely that punishers reason explicitly in this way; indeed, past work suggests punishers are typically motivated by retribution (Carlsmith, Darley, & Robinson, 2002; Darley, Carlsmith, & Robinson, 2000) rather than more sophisticated concerns about correcting behaviour. Rather, this perspective suggests that our underlying, automatic retributive impulses are structured in such a way that they adaptively motivate us to punish when it will deter future harm to us, i. e., in cases of negative outcomes. For now, we can only speculate; further research is needed on the reason, or reasons, that punishment is influenced by outcomes more than other types of moral judgments.

3.8. Towards an integrated model of moral judgment

While we have focused on differences between the models and accounts discussed here, they are not necessarily mutually exclusive, and may in several ways be complementary. In particular, while the current findings regarding blame and wrongness judgments of accidental agents seem to be well-explained by the negligence-based account, it does not explain the direct outcome effect on punishment judgments. This might be better accounted for by a model that places greater emphasis on the influence of outcomes. For example, Cushman and colleagues' (2008; 2013) two-process model might be extended to include a focus on unintended, though potentially foreseeable, negative consequences, and a role for negligence as a substantial influence on moral judgments. Our findings suggest that the mental state process, or processes, are triggered not only by the beliefs that agents *have* about possible consequences of their actions, but also by the beliefs that they *should* have: for example, Garmess is judged to be punishable, blameworthy and wrong not because he believed that the person upstairs was a burglar (which is

reasonable), but because he should have realised, first, that he might end up killing someone, and second, that it might not be a burglar but an innocent friend. Importantly, the processes that are sensitive to the causes of outcomes and to negligence are not merely additive, but interact so that punishment is assigned only if both processes are triggered. Moreover, participants' re-evaluation of the agent's beliefs when a negative outcome occurs (e.g., Garmess' belief that Drake was a burglar was re-assessed when it led to the death of an innocent friend) suggests a connection between the causally-sensitive and mental state processes, by which information about one can change how the other is evaluated.[10]

Another example of the potential for integration of apparently contradictory approaches within a single, holistic model arises from the possibility that judgments of accidental agents are in certain settings more outcome-based and influenced by intuitions, emotions, cognitive biases and motivations, and in others more negligence-based and influenced by systematic updating and the rational application of moral principles. For instance, Gino et al. (2008) reported that, when participants were asked to take an "intuitive" perspective based on gut feelings, there was more evidence of an outcome bias than when they took a "deliberative" perspective, from which they made more rational and objective judgments. Similarly, Martin, Buon, and Cushman (2021) report that judgments tend to be more outcome-based under cognitive load. It is also possible that influences between perceived negligence and judgments are bidirectional. For example, in some contexts, people's initial assessments of negligence might influence their judgments, which in turn lead them to re-evaluate their negligence attributions.

3.9. Implications for policy and practice

These findings suggest that people who judge accidental agents such as defendants, employees and children, are unlikely to punish when – by chance – the outcome is neutral, but that they are likely to punish when the same action, which is equally negligent, blameworthy and wrong, results in a negative outcome. To counter this moral, legal and social injustice that occurs as a result of moral luck, the current findings suggest that people should be encouraged to consider how negligent, blameworthy and wrong accidental agents are before assigning them punishment: these judgments are much less sensitive to outcome effects and moral luck, and therefore are more rational and fair. The finding of imbalance in moral luck indicates that this approach is likely to be more effective in cases where the agent has luckily not caused harm than when they unluckily have caused harm. This approach could be integrated with methods that recent studies have shown to reduce moral luck and therefore improve the justice of judgments, such as giving people both positive and negative outcomes, asking them to consider counterfactuals, or taking a deliberative rather than intuitive perspective (Gino et al., 2008; Kneer & Machery, 2019; Kneer & Skoczeń, 2023; Lench et al., 2015).

3.10. Limitations and future research

For several reasons we urge caution in interpreting these results. One is that we used only the Garmess vignettes in these experiments, and it is possible that responses to different stories would be different. However, the main findings of Experiment 2 were similar to those of Nobes and Martin (2022), in which participants were told about a driver steering through a leaf-pile. Moreover, Kneer and Machery (2019) and Kneer and Skoczeń (2023) both used two other sets of vignettes, and several of their main findings – in particular the principal role of negligence in moral judgments, and the additional impact of outcome on punishment judgments – are consistent with ours. Margoni et al. (2019) used four different scenarios, and Margoni & Surian, 2021 three more, and they, too, reported strong negligence effects on moral judgments. While further research using a wider range of stimuli is required, the available evidence from studies that have manipulated negligence therefore indicates that the key findings generalize to very different sets of vignettes.

A similar point concerns the use of weapon type to manipulate negligence. Doing so allowed us to replicate and build on Enzle and Hawkins' (1992) and Mazzocco et al.'s (2004) studies, but there are several possible problems. For example, the distinction between Garmess having a gun, bat, or no weapon might be interpreted by some participants as indicating more about his intention (or lack of intention) to kill, or his personal character as a weapon-keeper and -user (or non-keeper and -user), than about his negligence, in which case their judgments of Garmess would reflect the influence of these other factors rather than of negligence. However, Enzle and Hawkins (1992), Mazzocco et al. (2004) and we in both our experiments found that participants did rate use of the gun to be more negligent than use of the bat, and so the manipulation has been well-validated.

Similarly, the low-negligence condition, in which Garmess merely called out, might be questioned on the grounds that it differed from the bat and gun conditions not only in terms of negligence, but also of intentionality and causal responsibility. However, our data indicate that these three conditions successfully created a clear contrast in perceived negligence, and that this explained much more of the variance in judgments than did causal responsibility or perceived intention. This indicates that, even if participants considered Garmess' actions less intentional and causally responsible than when he used a weapon, it was perceived negligence, rather than these other factors, that was the primary influence on their judgments. In addition, the similarities between our findings and those of studies in which negligence was manipulated more directly (e.g., Margoni et al., 2019; Nobes & Martin, 2022) support the view that the weapon-type works as a valid negligence IV, and strengthens the case that the findings of this body of research are generalizable across varying stimuli and methods of manipulating the independent variables. Nonetheless, future researchers are encouraged to use more direct methods to manipulate negligence by, for example, explicitly stating that, and how, the agent was or was not negligent.

Another possible criticism is that in our two experiments the participants were predominantly well-educated and British or American. At this stage we cannot know the extent to which our findings generalize to other populations. Similarly, most were young adults. It has been suggested that young children's (Nobes et al., 2009; Nobes et al., 2017) and older adults' (Margoni et al., 2019) moral judgments appear to be more outcome-based than younger adults' because they tend to overattribute negligence, that is, they assume that accidental harm-doers are more negligent. If so, we might expect perceived negligence in these age groups to be an even stronger influence on moral judgments than is reported here.

Future research might investigate the proposal that, at least regarding blame and wrongness judgments, moral luck occurs because negative outcomes lead people to update, or re-evaluate, the riskiness, and hence negligence, of actions. This could be tested by using a similar approach to Monroe and Malle (2019) in which participants assessed agents' negligence and culpability before and after the actual negligence

---

[10] Kneer and Machery (2019) and Kneer and Skoczeń (2023) argue that their studies confirm Cushman's (2008) two-process model, but in our view their findings are largely consistent with ours, and lend considerably more support to the negligence-based account. According to the model, one of the two processes is sensitive to mental states but, whereas Cushman and colleagues refer almost exclusively to intentions, desires and beliefs, Kneer and colleagues consider the same process to be sensitive to negligence. However, negligence and intention are independent factors – a well-intentioned agent can be negligent or non-negligent – and so it is difficult to see how a single mental state process could be sensitive to both. Treating negligence and intention as separate and independent influences leads to very different predictions and explanations of judgments of accidental agents from those based on the two-process model.

and outcome information was presented. It could be combined with Kneer and Skoczeń's (2023) methods in which each participant was given one (between-subjects) or both (within-subjects and counterfactual designs) outcomes. If the updating explanation were correct, assessments of probability, negligence and judgments would be updated systematically as information became available, rather than influenced primarily by the actual outcome, as the motivated-blame and hindsight bias approaches would predict.

4. Summary

Regarding blame and wrongness, the results of the experiments reported here are consistent with the view that negligence is the principal influence on judgments of accidental actions, and should be considered one of the "major building blocks of moral judgment" (Kurdi et al., 2020, p. 2). They also indicate that outcome effects on these judgments are largely explained by the differential attribution of negligence: accidental agents who unluckily cause harm are attributed more negligence than those who luckily cause no harm. This might be accounted for by a hindsight bias (e.g., Kneer & Machery, 2019; Margoni et al., 2023), and we have proposed an alternative explanation, similar to Monroe and Malle's (2019), according to which outcomes provide information that leads participants to systematically update their assessments of agents' negligence and culpability.

Punishment judgments are strongly influenced by negligence too, but also by outcome; agents who are considered negligent, and hence blameworthy and wrong, are assigned punishment only if the outcome is negative. This is only partially explained by the negligence-based account, and there is considerable potential for developing a model that emphasizes the key roles of both negligence and outcome in moral judgments of accidental agents.

These findings corroborate and expand on those of other studies that have manipulated and/or measured the role of negligence in moral judgments of accidental agents (e.g., Kneer & Machery, 2019; Margoni et al., 2019; Nobes et al., 2009; Nobes & Martin, 2022). They indicate that previous reports of strong outcome effects – particularly on blame and wrongness judgments – are often better explained by the differential attribution of negligence. There is now a strong body of evidence that indicates that in most studies in this area the influence of negligence on moral judgments has been underestimated, and that of outcome has been overestimated.

Data availability

Data are available at https://osf.io/zm9bd/?view_only=4bd1634b76ea4794a64c5ae4e1e1c25d

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2023.104456.

References

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126*(4), 556–574. https://doi.org/10.1037/0033-2909.126.4.556

Alicke, M. D. (2014). Evaluating blame hypotheses. *Psychological Inquiry, 25*(2), 187–192. https://doi.org/10.1080/1047840X.2014.902723

Alicke, M. D., & Davis, T. L. (1989). The role of a posteriori victim information in judgments of blame and sanction. *Journal of Experimental Social Psychology, 25*(4), 362–377. https://doi.org/10.1016/0022-1031(89)90028-0

Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M., Fitzpatrick, S., Gurven, M., … Scelza, B. A. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences, 113*(17), 4688–4693. https://doi.org/10.1073/pnas.1522070113

Bernstein, D. M., Erdfelder, E., Meltzoff, A. N., Peria, W., & Loftus, G. R. (2011). Hindsight bias from 3 to 95 years of age. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 378–391. https://doi.org/10.1037/a0021971

Bornstein, B. H. (1998). From compassion to compensation: The effect of injury severity on mock jurors' liability judgments. *Journal of Applied Social Psychology, 28*(16), 1477–1502. https://doi.org/10.1111/j.1559-1816.1998.tb01687.x

Carlsmith, K., Darley, J., & Robinson, P. (2002). Why do we punish?: Deterrence and just deserts as motives for punishment. *Journal of Personality and Social Psychology, 83*(2), 284–299. https://doi.org/10.1037/0022-3514.83.2.284

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. (2015). Deconstructing intent to reconstruct morality. *Current Opinion in Psychology, 6*, 97–103. https://doi.org/10.1016/j.copsyc.2015.06.003

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

Darley, J. M., Carlsmith, K. M., & Robinson, P. H. (2000). Incapacitation and just deserts as motives for punishment. *Law and Human Behavior, 24*(6), 659–683. https://doi.org/10.1023/A:1005552203727

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

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. (2008). No harm, no foul: The outcome bias in ethical judgments (Harvard Business School working paper, Issue 08–080). 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. D. (2008). The secret joke of Kant's soul. In W. Sinnott-Armstrong (Ed.), *Vol. 3. The neuroscience of morality: Emotion, brain disorders, and development* (pp. 35–80). MIT Press.

Gunpolicy.org. (2020). Gun facts, figures and the law. https://www.gunpolicy.org.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review, 108*, 814–834. https://doi.org/10.1037/0033-295X.108.4.814

Hartman, R. (2017). *In defense of moral luck: Why luck often affects praiseworthiness and blameworthiness*. Routledge. doi:113829344X.

Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin, 107*, 311–327.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2nd ed.). Guilford Press.

Kant, I. (1784/1998). *Groundwork of the metaphysics of morals (M. Gregor, Ed. & Trans.)*. Cambridge University Press.

Kim, M., Park, B., & Young, L. (2020). The psychology of motivated versus rational impression updating. *Trends in Cognitive Sciences, 24*, 101–111. https://doi.org/10.1016/j.tics.2019.12.001

Kneer, M., & Bourgeois-Gironde, S. (2017). Mens Rea ascription, expertise and outcome effects: Professional judges surveyed. *Cognition, 169*, 139–146. https://doi.org/10.1016/j.cognition.2017.08.008

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., & Skoczeń, I. (2023). Outcome effects, moral luck and the hindsight bias. *Cognition, 232*, Article 105258. https://doi.org/10.1016/j.cognition.2022.105258

Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis, 63*, 190–194. https://www.jstor.org/stable/3329308.

Kurdi, B., Krosch, A. R., & Ferguson, M. J. (2020). Implicit evaluations of moral agents reflect intent and outcome. *Journal of Experimental Social Psychology, 90*, 1–12. https://doi.org/10.1016/j.jesp.2020.103990

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., Domsky, D., 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

Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry, 25*(2), 147–186. https://doi.org/10.1080/1047840X.2014.877340

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., Geipel, J., Hadjichristidis, C., & Surian, L. (2023). Adult age-related differences in moral judgment: The role of hindsight. https://osf.io/ugct6.

Margoni, F., & Surian, L. (2021). Judging accidental harm: Due care and foreseeability of side effects. *Current Psychology, 41*, 8774–8783. https://doi.org/10.1007/s12144-020-01334-7

Martin, J. W., Buon, M., & Cushman, F. (2021). The effect of cognitive load on intent-based moral judgment. *Cognitive Science, 45*, 1–25. Article e12965 https://doi.org/10.1111/cogs.12965.

Martin, J. W., & Cushman, F. (2015). To punish or to leave: Distinct cognitive processes underlie partner control and partner choice behaviors. *PLoS One, 10*(4). https://doi.org/10.1371/journal.pone.0125193

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

McNamara, R. A., Willard, A. K., Norenzayan, A., & Henrich, J. (2019). Weighing outcome vs. intent across societies: How cultural models of mind shape moral reasoning. *Cognition, 182*, 95–108. https://doi.org/10.1016/j.cognition.2018.09.008

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

Nagel, T. (1979). *Mortal questions.* Cambridge University Press.

National Center for Injury Prevention and Control. (2019). Fatal Injury Reports, National, Regional and State, 1981–2018. https://webappa.cdc.gov/sasweb/ncipc/mortrate.html.

Nelkin, D. K. (2021). Moral luck. *The Stanford Encyclopedia of Philosophy.* https://plato.stanford.edu/archives/sum2021/entries/moral-luck/.

Nobes, G., & Martin, J. W. (2022). They should have known better: The roles of negligence and outcome in moral judgements of accidental actions. *British Journal of Psychology, 113*, 370–395. https://doi.org/10.1111/bjop.12536

Nobes, G., Panagiotaki, G., & Bartholomew, K. (2016). The influence of intention, outcome and question-wording on children's and adults' moral judgments. *Cognition, 157*, 190–204. https://doi.org/10.1016/j.cognition.2016.08.019

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

Office for National Statistics. (2020). Appendix tables: Homicide in England and Wales. https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/datasets/appendixtableshomicideinenglandandwales.

Patil, I., Calò, M., Fornasier, F., Young, L., & Silani, G. (2017). Neuroanatomical correlates of forgiving unintentional harms. *Scientific Reports, 7*. https://doi.org/10.1038/srep45967

Perugini, M., Gallucci, M., & Costantini, G. (2018). A practical primer to power analysis for simple experimental designs. *International Review of Social Psychology, 31*. https://doi.org/10.5334/irsp.181. Article 20.

Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind & Language, 24*, 586–604. https://doi.org/10.1111/j.1468-0017.2009.01375.x

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

Rosset, E. (2008). It's no accident: Our bias for intentional explanations. *Cognition, 108*, 771–780. https://doi.org/10.1016/j.cognition.2008.07.001

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. https://doi.org.uea.idm.oclc.org/10.1037/h0080138.

Siegal, M., & Peterson, C. C. (1998). Preschoolers' understanding of lies and innocent and negligent mistakes. *Developmental Psychology, 34*(2), 332–341. https://doi.org/10.1037/0012-1649.34.2.332

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