



Judging accidental harm: Due care and foreseeability of side effects

Francesco Margoni^{1,2} · Luca Surian¹

Accepted: 30 December 2020 / Published online: 17 January 2021
© The Author(s) 2021

Abstract

Both in philosophy and in cognitive psychology, models of moral judgment posit that individuals take into account both agents' intentions and actions' outcomes. The present research focused on a third crucial piece of information, agents' negligence. In Study 1, participants judged the moral wrongness and punishability of agents' actions that resulted in negative side effects. In the scenarios, we orthogonally manipulated whether the agent acted with or without due care and whether she had or did not have information to foresee the negative side effects of her actions. Participants judged careless agents more condemnable than careful agents, especially when negative side effects could have been easily foreseen. In Study 2, we manipulated due care in acting in cases where the agent's primary intention was to bring about a certain outcome without knowing that such outcome would actually be harmful. Here information about the foreseeability of negative outcomes was not provided, and participants judged actions performed with care more wrong and punishable than actions performed without care. This suggests that sometimes acting carefully and nevertheless causing harm may constitute evidence of the presence of negative intentions in the agents or evidence of the fact that agents indeed could have foreseen the negative effects of their actions. Together, these findings indicate that carefulness in acting and foreseeability are highly intertwined in moral judgment, and highlight the need to improve existing processing models of moral judgment to account for people's evaluation of agents and actions whenever negligence can be attributed.

Keywords Ethics · Moral judgment · Intentions · Negligence · Carefulness · Foreseeability

A long-lasting tradition in moral philosophy and jurisprudence emphasizes the relevance of information about agents' intentions for moral judgment (Abelard, XII century/1971; Anscombe, 1957; Kant, 1785/1959; Williams, 1953). However, while philosophical work emphasizes the role of intention, it does not say much about how the human mind processes this information. Aiming at investigating information processing in the human mind, a now large body of literature in moral psychology has revealed that in their moral judgments people tend to assess both agents' causal responsibility for negative or positive outcomes and agents' intentions, and often rely mostly on intentions (e.g., Barrett et al., 2016; Cushman, 2008, 2015; Malle & Knobe, 1997; Margoni, Geipel, Hadjichristidis, & Surian, 2018; Young & Saxe, 2009a).

It has indeed been found that even infants hold sociomoral expectations and express intuitive preferences that take into account agents' intentions and prioritize them over actions' outcomes (Dunfield & Kuhlmeier, 2010; Hamlin, 2013; Hamlin, Ullman, Tenenbaum, Goodman, & Baker, 2013; Strid & Meristo, 2020; Woo, Steckler, Le, & Hamlin, 2017). Furthermore, we know that at least by five years of age, children consistently weigh intentions more than outcomes when they are asked to express a moral judgment in tasks that are more demanding compared to those employed with infants (Cushman, Sheketoff, Wharton, & Carey, 2013; Margoni & Surian, 2017, 2020; Proft & Rakoczy, 2018; see Margoni & Surian, 2016 for a review).

A number of processing models of moral judgment have been proposed in order to account for the relative weight of intentions and causal responsibility for outcomes. In a seminal work, Heider (1958) proposed that people first analyze the causal responsibility for harm, and only subsequently analyze internal or mental state factors such as intentions (Darley & Shultz, 1990; Weiner, 1995). Importantly, according to this model, (a) both outcome and intention are needed to attribute moral responsibility, blame and punishability; and (b) causal and intentional information are integrated by a single cognitive process. So, to give an example, the judgment that the

✉ Francesco Margoni
francesco.margoni@gmail.com

¹ Department of Psychology and Cognitive Sciences, University of Trento, Trento, Italy

² Department of Psychology, University of Oslo, Oslo, Norway

agent deserves to be punished has been described as the outcome of a process that first assesses the causal responsibility of the agent's action for a certain negative outcome, and then attributes *moral* responsibility to the agent (Shultz, Schleifer, & Altman, 1981; Shultz, Wright, & Schleifer, 1986).

More recently, Fiery Cushman claimed that outcomes and intentions are assessed by two distinct and independent cognitive processes (Cushman, 2015; Cushman et al., 2013). While one of these processes focuses on the causal responsibility for harm, the other focuses on the agent's mental states. This two-process model predicts that sometimes a conflict will arise between incompatible responses, for example in those cases in which the agent accidentally harms someone (bad outcome, no negative intention) or intends to harm someone but then fails (bad intention, no negative outcome).

Cases of *accidental harm* are particularly demanding to evaluate because most likely people have to inhibit a prepotent outcome-based response (a salient bad outcome occurred and the agent is causally responsible for it) and select an intent-based response (the agent did not have a bad intention, and thus is not morally condemnable) (Margoni & Surian, 2016, 2020; Young & Saxe, 2009b). The task of inhibiting and selecting responses may be further complicated by the fact that we can attribute *negligence* to accidental transgressors ('perhaps they should have been more careful and/or they should have known better'). A spontaneous tendency to attribute negligence to accidental wrongdoers has indeed been found to be particularly strong in preschoolers (Nobes, Panagiotaki, & Pawson, 2009; Nobes, Panagiotaki, & Engelhardt, 2017), as well as in old adults (Margoni, Geipel, Hadjichristidis, & Surian, 2019), possibly because of developmental changes in inhibitory control abilities (i.e. increase during childhood and decrease in the old age; Buon, Seara-Cardoso, & Viding, 2016; Margoni & Surian, 2016, 2020). Another source of difficulty may be the human proclivity to think that the controlling forces behind harmful consequences are people rather than the environment, where a possible rationale for this tendency is the fact that human actions are easier to rectify compared to environmental events (Alicke, 2000; see also Alicke, Buckingham, Zell, & Davis, 2008; Alicke, Rose, & Bloom, 2011).

In general, using accidental harm scenarios helps researchers to assess whether people focus not simply on intentions and outcomes but also on negligence and on whether the agents could have had control over the outcomes of their actions. However, what is negligence? Let us start with a simple example: a driver that, because of insufficient concentration on the road (e.g., he was sending a text message), caused an accident and killed someone. Despite the fact that the action of killing was unplanned, the driver would likely be judged as morally condemnable. This is because we would attribute 'negligence', that is, we would judge that the driver did not pay enough attention to the road (he was careless) where however he could have easily known and should have known that

not paying attention when driving can cause serious harm (D'Arcy, 1963; Hart, 1968; Shen, Hoffman, Jones, & Greene, 2011). Negligence thus contributes to the *mens rea* account of culpability, as it refers to an aspect of the agent's mental state (i.e. we could reasonably expect that he should have known better), but still it has to be differentiated from intention.

According to this definition, negligence would be attributed when the agent acted without due care (on the physical behavioral side, he was sending text messages and not paying attention to the road) *and* he could have predicted or foreseen the negative side effects of his behavior (on the cognitive side, he likely knew or should have known that being distracted not so rarely can result in an accident). Moreover, it is also worth noting that negligence is often attributed to agents whose harmful outcomes were unintended, that is, the agent wanted to text a message but did not intend to kill someone (that was a negative and unintended side effect of his action).

Negligence may thus be a relevant input in a computational model of ordinary people's moral judgment (Weiner, 1995). Research has indeed shown that we begin to distinguish between innocent accidental harm and accidental harm caused by a negligent agent already in preschool (Schleifer, Shultz, & Lefebvre-Pinard, 1983; Siegal & Peterson, 1998), and at least by middle childhood we can fully use negligence information to assign moral responsibility (Shultz & Wright, 1985; see also Nelson-le Gall, 1985). Research on adult moral judgment suggests that, similarly to what happens in a legal system, we often seek to determine what risks the agent could have reasonably foreseen, and our ratings of responsibility decrease depending on the number of precautions that the agent took to avoid potential harmful consequences (Karlovac & Darley, 1988; see also Enzle & Hawkins, 1992; Finkel & Groscup, 1997). Moreover, a recent study showed that both children and adults attend to transgressor's negligence but, interestingly enough, also to victim's negligence (Mulvey, Gönültaş, & Richardson, 2020).

The Present Research

The goal of the present research was to investigate how adults attend to both agents' carefulness in acting (at the physical behavioral level) and foreseeability of the negative side effects of their actions in evaluating cases of accidental harm. To this aim, in Study 1 we asked participants to morally evaluate unintended negative outcome scenarios that were generated by varying the agent's carefulness in acting and the foreseeability of the action's outcomes. More specifically, we asked participants to evaluate the case of Helen, a waitress who either knowing or not knowing that under a table there was a cradle containing a baby, accidentally tramples on the cradle

hitting the baby while moving either carefully or carelessly among the tables.

According to recent research, the information processing of moral cases may follow two distinct paths. If the transgression is judged to be intentional, people consider the agent's reasons or motives for acting when computing blame. If instead the transgression is judged to be unintentional, then people consider how preventable the transgression was (Malle, Guglielmo, & Monroe, 2014; Monroe & Malle, 2017; see also Lagnado & Channon, 2008). Therefore, we predicted that participants in our study would weigh the information about care in acting more when the agents could have easily prevented the negative side effects of their action than when the agents, given the circumstances, could not have easily prevented the negative side effects. That is, we predicted that participants would condemn the agent more when she moved carelessly *and* could have easily prevented the harmful outcome because she knew about the presence of the baby.

In the scenarios used in Experiment 1, the agents always brought about an accidental harm by performing an accidental movement (tripping over the cradle). In Experiment 2, we looked at whether participants respond similarly when they are told a story about an agent that acts purposefully to bring about a certain outcome but, unbeknownst to him, his action violated a moral norm. Can this scenario reverse the effect of care we predicted to be found in Experiment 1? Would people condemn such an accidental transgressor more when he acted carefully instead of carelessly? Suppose that a certain outcome is the primary goal of the agent who, however, lacks some relevant pieces of information that would allow him to realize that his action constitutes a moral transgression (due to its negative side effect). Then imagine that the agent performs his action being very attentive (i.e. with care) or, instead, absentmindedly (i.e. without care). For example, let us consider the case of a young man who wants to take the last free seat on the bus not knowing that an old man is also trying to find a seat, and he does it either with or without care (e.g. by rushing to make sure to arrive first vs. by performing his action absentmindedly). Here a possibility is that participants would see carefulness not as a mitigating factor reducing the blameworthiness of the agent, but instead as signaling that the purposeful action (taking the last seat) was performed with some awareness of its negative consequences (the old man is forced to travel standing).

Study 2 aimed at addressing this possibility by asking participants to judge cases like the one we gave as an example where, unlike in Study 1 (in which Helen did not purposefully trip over the cradle), the agent performed an action to bring about a certain outcome (i.e. the young man indeed wanted to occupy the last free seat), although he was not aware of the fact that his action would cause harm. Finding here that careful accidental transgressors are condemned more than careless accidental transgressors will suggest that carefulness can be

an aggravating factor when it applies on the performance of a purposeful action (i.e. the young man wanted to occupy the last free seat), because in this latter case it can be interpreted as a cue of a negative intention or that the agent actually foresaw the negative side effects of his action. Taken together, Study 1 and Study 2 aimed at showing how both carefulness in performing some actions and foreseeability of the negative side effects are highly intertwined elements that can be central in the evaluation of accidental harm.

Study 1

Study 1 orthogonally manipulated agent's carefulness (i.e. whether the agent acted with care) and foreseeability of the action's side effects (i.e. the agent's knowledge state, whether she knew about the dangerousness of her action) in a case of accidental harm. Participants judged the resulting scenarios by rating how much the agent's behavior is morally wrong, how much the agent deserves punishment, and how much responsibility the agent has.

Method

Participants

Participants were 48 adults (36 female, $M_{age} = 21.67$ years, $SD = 1.65$), recruited among students enrolled in psychology courses at the University of Trento. All participants provided written and informed consent, and the local University Ethics Committee approved the experimental procedure. The sample size was determined by an a priori power analysis for a mixed-factor 2 (*care*: present, absent) \times 2 (*foreseeability*: present, absent) ANOVA. To detect as statistically significant a medium effect size $f = 0.25$ (based on prior research on negligence, Margoni et al., 2019a), with alpha set at 0.05, a power of 0.80, correlations among measurements = 0.50, a minimum total sample size of 34 participants was required. However, we recruited a slightly larger sample of 48 participants as a safeguard measure to protect against the possibility of the unknown true effect being smaller than 0.25. In particular, a 2×2 mixed-factor ANOVA with 48 participants and power set at 0.80 can detect an effect as small as $f = 0.20$.

Materials and Procedure

Participants were tested in groups. They completed a paper-and-pen questionnaire. We adopted a 2×2 design, manipulating *care* (agent acted with or without due care) as a within-subjects factor and *foreseeability* (agent knew or did not know about the dangerousness of her action) as a between-subjects factor. We thus generated four stories about Helen, a waitress who accidentally tramples on a cradle containing a baby.

Helen is either informed or not informed about the presence of the baby and its position, and moves around either with or without care. An example of the stories used (where the agent acts with care and harm is easily foreseeable) is the following (but see Appendix 1 for the complete set):

Helen is a waitress waiting tables in a classy mountain shelter. Under a table occupied by some tourists, a newborn is sleeping in a cradle. Helen knows that under the table there is a baby because she saw it; Helen just wants to wait tables. Helen moves carefully among the tables. However, Helen steps accidentally on the cradle and hits the baby.

Participants answered the following three questions with a seven-point scale (1 = none; 4 = some; 7 = very much):

- *Wrongness question*: “How morally wrong is Helen’s behavior?”
- *Punishability question*: “How much does Helen deserve to be punished?”
- *Responsibility question*: “How much responsibility does Helen have for what happened?”

Each participant judged only two stories (either with a foreseeable side effect or without it). We counterbalanced questions order across participants using a Latin Square.

Results and Discussion

We analyzed participants’ responses with three separate 2 (Care: present vs. absent) \times 2 (Foreseeability: present vs. absent) mixed-factor ANOVA, one for each judgment type (wrongness, punishability, responsibility). For each judgment type, the analysis revealed a main effect of care, $F(1, 46) \geq 19.07$, $ps < .001$, $f \geq 0.64$ (on average, participants gave higher ratings when judging careless actions), a main effect of foreseeability, $F(1, 46) \geq 12.20$, $ps \leq .001$, $f \geq 0.51$ (on average, participants gave higher ratings when judging side effects that were easily foreseeable), and a significant interaction between the two factors, $F(1, 46) \geq 5.17$, $ps \leq .028$, $f \geq 0.33$ (see Fig. 1 and Table 1).

Planned comparisons revealed that participants judged the careless action significantly more wrong than the careful action when the negative side effect of the action was easily foreseeable, $F(1, 46) = 57.70$, $p < .001$, Cohen’s $d = 0.98$, but did not do so to the same extent when the side effect was not so easily foreseeable, $F(1, 46) = 9.01$, $p = .004$, $d = 0.53$ (see Table 2 for means and SDs). Indeed, the effect size was halved, from $d = 0.98$ to $d = 0.53$. Analyses also revealed that when the negative side effect was foreseeable, participants judged the careless agent more punishable than the careful

agent, $F(1, 46) = 65.12$, $p < .001$, $d = 1.06$, but did not do so when the side effect was described as not so easily foreseeable, $F(1, 46) = 0.26$, $p = .612$, $d = 0.10$. Lastly, participants judged the careless agent more responsible for the negative outcome than the careful agent when the outcome was foreseeable, $F(1, 46) = 74.41$, $p < .001$, $d = 1.25$, but did not do so to the same extent when the outcome was not easily foreseeable, $F(1, 46) = 13.14$, $p = .001$, $d = 0.58$. Indeed, the effect size was halved, from $d = 1.25$ to $d = 0.58$. These results suggest that the effect of care is larger when the side effect could have been easily foreseen.

Moreover, it may be noted that mean scores for both wrongness and punishability judgments ranged from about 1.5 to about 4 (see Table 2), thus ranged from *none* to *some*, confirming prior work (e.g., Cushman, 2008; Margoni et al., 2018) suggesting that adults’ moral judgments rely more on intention (in this case absent) than on outcome.

In sum, Study 1 showed that although participants’ judgments consistently relied on the information about due care in acting, this piece of information was particularly crucial when the harmful side effect of the agent’s action was easily foreseeable given the knowledge held by the agent.

Study 2

Study 2 complemented Study 1 by investigating whether the effect of care found in Study 1 can be reversed, that is, whether under certain circumstances people would judge accidental transgressors who acted with care more condemnable than accidental transgressors who acted without care. Participants in Study 2 evaluated cases in which the action or set of movements that eventually resulted in a harmful outcome were intended by the agent who however lacked relevant pieces of information to see his action as harmful, and acted either being attentive or being absentminded (carefulness of the agent’s movements was thus manipulated). We reasoned that in judging these cases, participants would associate acting carefully with the agent having in reality a negative intention or at least having foreseen the negative side effects of his action, whereas they would associate acting absentmindedly with being simply distracted and not having any real negative intention or clue about the negative side effects.

Method

Participants

Participants were 30 adults (18 female, $M_{age} = 25.96$ years, $SD = 3.71$), recruited like in Study 1 or from the urban middle-class area surrounding the campus. All participants provided informed consent. The sample size was determined by an a priori power analysis for a two-tailed paired samples t

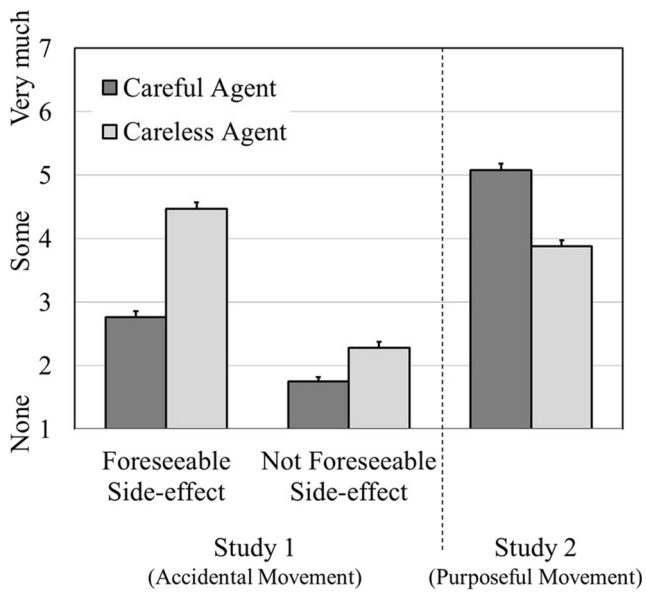


Fig. 1 Mean combined judgments (wrongness, punishability, responsibility) of an accidental harm produced by an accidental movement (Study 1) or by a purposeful movement (Study 2). Error bars show the magnitude of the standard error of the mean

test. To detect an effect size $d = 0.53$ (based on the smaller effect detected in Study 1), with alpha set at 0.05 and a power of 0.80, a minimum sample size of 30 participants was required.

Materials and Procedure

Participants completed a paper-and-pen questionnaire in a quiet room, at their house or in the laboratory, and evaluated the wrongness and the punishability of accidental harmful actions that were performed either with or without care (see Study 1 for the question-wording). We had two contexts (*injustice* and *malevolence*) and thus four different stories. The contexts were inspired by two classical virtues of the Latin and Greek conceptual world (Marcus Aurelius, II century AD/2006): *iustitia*, manifested by agents that act fairly, and *benevolentia*, manifested by agents that act to promote others' well-being.

In the injustice context, the boss delivers to Jack both his and his colleague's wage thinking that the two are old friends. However, Jack does not know this. He thinks that the entire

sum of money is just for him, and takes the money either carefully or absentmindedly. In the malevolence context, James is about to get on the bus and take a seat. An old man is getting on too. James does not have any particular evil intention. He just occupies the last free seat either without any concerns for arriving first or making sure to sit down before everyone else. Scenarios were presented in a randomized order, and questions order was counterbalanced across participants. An example of the stories used (injustice context, careful action) is the following (see Appendix 2 for the complete set):

During the summer, Jack and Thomas, who are strangers, work a few days together for a company that organizes summer camps. During the distribution of wages, the boss delivers all the money to Jack as Thomas is sick and the boss thinks that the two employees are old friends. Jack wants to keep all the money for himself, because he erroneously thinks that the wage is just for him. Jack takes the money carefully. Eventually, Jack keeps also Thomas' money.

Results and Discussion

In the following analyses, responses to the two contexts were pooled as we focused our analysis on the influence of care on participants' evaluations. Participants judged accidental harmful actions performed with care more wrong and punishable than accidental harmful actions performed without care, $t(29) \geq 3.28$, $ps \leq .003$ (see Fig. 1 and Table 2). This may suggest that participants inferred that, compared to careless agents, careful agents were more intentioned or determined in bringing about the negative outcome or at least were more aware of the eventual negative side effect of their actions. Unlike when the action that causes harm is unplanned (i.e. stumbling on a cradle and hitting a baby or killing someone because texting while driving), when the action that accidentally brings about negative consequences is intended (i.e. taking all the money), the fact that the agents acted carefully may signal that, at least to some extent, they intended to harm or indeed could have foreseen the side effects of their actions.

Table 1 Main Results of the Mixed Factor ANOVAs for Wrongness, Punishability and Responsibility Judgments in Study 1

	Wrongness			Punishability			Responsibility		
	$F(1, 46)$	p	η_p^2	$F(1, 46)$	p	η_p^2	$F(1, 46)$	p	η_p^2
Care	28.17	< .001	.380	19.07	< .001	.293	37.53	< .001	.449
Foreseeability	12.20	.001	.210	22.60	< .001	.329	29.94	< .001	.394
Care * Foreseeability	5.17	.028	.101	13.90	.001	.232	6.19	.016	.119

Table 2 Means and SDs of Participants' Judgments in Study 1 and 2

	Careful						Careless					
	Wrong		Punish		Responsible		Wrong		Punish		Responsible	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Study 1 Foreseeable	2.21	1.56	2.50	1.41	3.58	1.82	3.88	1.83	4.08	1.56	5.46	1.10
Study 1 Not foreseeable	1.42	0.93	1.63	1.06	2.21	1.35	2.08	1.50	1.75	1.33	3.00	1.35
Study 2	5.57	1.30	4.58	1.51	–	–	4.37	1.50	3.40	1.48	–	–

These spontaneous inferences may have led participants in Study 2 to condemn careful agents more.

An indirect way to test this hypothesis leans on previous work showing that wrongness judgments, compared to punishability judgments, are more influenced by intentions (Cushman, 2008, 2015). If this is the case, and if careful actions in Study 2 were indeed considered more guided by evil intentions than careless actions, we can expect the effect of care we reported to be bigger for wrongness judgments. However, contrary to this prediction, the interaction between care (present, absent) and type of judgment (wrongness, punishment) was not statistically significant, $F(1, 29) = 0.01$, $p = .950$.

General Discussion

In this research we aimed at assessing how the information about due care in acting influences moral judgment as a function of the foreseeability of the negative side effects of the action. In Study 1, we found a main effect of both due care and foreseeability. Unintended negative side effects caused by agents who acted without care were condemned more than those caused by careful agents, and unintended negative side effects caused by agents who could have easily predicted them were condemned more than those caused by agents who could not have easily predicted them. Importantly, we also found a statistically significant interaction between due care and foreseeability, indicating that the information about due care in acting weighed more when the agent could have prevented the negative side effect (e.g. she knew where the baby was). This suggests that agents are judged negligent and therefore morally condemnable especially when they are both careless and lacking in foresight (given that they had sufficient contextual information to prevent the damage).

In Study 2, we instead reported a case in which careful agents were condemned more than careless agents. Crucially, the scenarios in Study 2 always described an agent who intended to bring about the main outcome in question although not knowing that his action also had a damaging effect for others. Here a possibility is that participants interpreted carefulness as signaling that the agent indeed wanted to cause harm or

at least was somehow aware of the possible negative side effects of his action, and carelessness as signaling that the agent did not really intend to act in the way he did and that he was simply distracted. In other words, the absence of due care, where due care was applied to actions purposefully directed at a certain outcome, may have been interpreted as evidence that the agent likely had a bad intention or had relevant information to foresee better, and thus should be blamed. However, the current research did not test this possibility directly. Therefore, a valuable suggestion for future research would be to ask participants to evaluate also to what extent in their opinion the accidental transgressor intended to bring about the negative side effects of their action.

It is noteworthy to remind that whether and how people take into account the foreseeability of side effects when generating their judgments has been investigated in the literature on the so-called Knobe effect. In this literature, the focus is on how moral valence affects the attribution of intentionality to side effects that could have been foreseen, but were explicitly described as unintended (Knobe, 2003, 2004). These studies have repeatedly found that both adults and young children deem negative but not positive side effects as being intentionally caused (Leslie, Knobe, & Cohen, 2006; Nichols & Ulatowski, 2007; Pellizzoni, Siegal, & Surian, 2009). Moreover, this attribution of intentionality is constrained by information about the agent's knowledge and beliefs about whether the side effects were foreseeable (Pellizzoni, Giroto, & Surian, 2010; see also Beebe, 2013; Uttich & Lombrozo, 2010). Note however that whereas this literature is mainly about whether and how the attribution of intentionality is affected by moral considerations, in the present study we focused instead on how carefulness and foreseeability affect moral judgment.

"Ethically, intention is everything", claimed Jean Piaget in his seminal work on the development of moral judgment (Piaget, 1932). Consistently, recent models of moral information processing have highlighted the centrality of the information about agents' intentions. As we discussed earlier in the Introduction, Cushman et al. (2013) posited a mental-state based process that analyzes agents' desires, motives and intentions, and that often prevails over an outcome-based

process, determining intent-based moral judgments in people assessing moral scenarios. Monroe and Malle (2017) further posited that judgments of intentionality importantly split moral information processing into two distinct paths. If the transgression was intentional, then people evaluate the agents' reasons and motives for acting and likely blame the agent. If, however, the act of transgression was unintended, then people begin to assess other aspects of the situation such as the preventability of the transgression.

Thus, in a way, yes, 'intention is everything in moral judgment'. However, cases like accidental harm help us see that people process other pieces of information too in order to make sense of the agents' actions and their relevant underlying mental states. The present research showed that when people judge accidental harm, they take into account due care in acting and foreseeability, but, most importantly, they rely on due care especially when the action's negative side effects were easily foreseeable. However, carelessness does not always associate with a tendency to blame, as this was true only when the transgression or harmful outcome was interpreted as fully unintended.

This research provided only a glance on the complex interconnection between the several pieces of information that undergo scrutiny during the evaluation of an accidental harm scenario. Future research could fruitfully investigate all those instances where people seem to judge unintended harm by simply relying on outcome to see whether this is actually the case or people are instead attributing a lack of due care to the agent and/or culpability because the outcome could and should have been prevented. Some of these instances may include clinical populations such as autistic individuals who, according to a widespread view, show difficulties in acquiring mentalistic reasoning skills and generating intent-based moral judgments (Baron-Cohen, Tager-Flusberg, & Lombardo, 2013; Buon et al., 2013; Margoni, Guglielmetti, & Surian, 2019; Zalla & Leboyer, 2011). Another instance may be populations from non-WEIRD (Western, educated, industrialized, rich, and democratic) societies such as people from small-scale societies that, according to a research conducted on groups like the Hazda from Africa and the Yasawa from the Pacific Islands, are apparently reluctant to consider intentions in their moral judgments (Barrett et al., 2016).

Some important limitations of the current research also need to be acknowledged, as they can further indicate valuable directions for future studies. First, we manipulated foreseeability by stating that the protagonist of the story either knew or did not know about the danger: Helen the waitress either knew or did not know that a baby was resting under the table. We reasoned that tripping on the cradle and hitting the child would have been judged more blameworthy if the agent could have easily foreseen this eventuality, that is, if she knew about the child in the first place. However, it would also be relevant to take into account the crucial distinction between the actual

knowledge possessed by the agent and the fact that the agent should have known about the danger. Indeed, in assessing negligence we often not only judge what a person knew about relevant facts but also and importantly what is reasonable to expect they *should* have known. If an agent knew about the risks and consciously disregards them, we may attribute recklessness, whereas negligence is more likely attributed when agents did not know about the risks but they should have been aware of them (Shen et al., 2011; see also Kneer & Machery, 2019). Future research can develop scenarios in which all these distinct behavioral and cognitive elements of negligence are controlled for and their relative contribution on moral judgment is assessed.

A second limitation is that in the scenario employed in Study 1 we manipulated due care simply by stating that the agent either "moves carefully among the tables" or "moves carelessly among the tables", thus leaving ample room for different interpretations of the words 'carefully' and 'carelessly'. In particular, at least to some extent, participants could have interpreted the word 'carefully' as meaning that the agent also moved with foresight, accurately considering the possible dangers. In order to avoid confounding care and foresight, future studies may want to employ scenarios where more information about care is provided, and perhaps may also want to directly ask participants how they interpreted such piece of information, as well as how careless they thought agents were.

Concluding Remarks

In sum, we reported evidence consistent with the view that due care, foreseeability of the negative consequences of actions, and intentionality of the action, all interact with each other in predictable manner when people process a moral scenario. In conclusion, we would like to note that negligence has often a crucial role in causing serious damage in our society, for example on the road or in the work environment (Reamer & Racette, 2015). Therefore, research on negligence and moral judgment may have potential far-reaching practical implications, as in planning and carrying out effective intervention programs aimed at reducing risk in daily life activities.

Funding Open Access funding provided by University of Oslo (incl Oslo University Hospital).

Data Availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Compliance with Ethical Standards

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Appendix 1 - Moral Scenario of Study 1

Scenarios were administered in Italian. Here below we provide the English translation complemented with the original version in Italian in square brackets.

Background: Helen is a waitress waiting tables in a classy mountain shelter. Under a table occupied by some tourists, a newborn is sleeping in a cradle. [Elena è una cameriera e sta servendo ai tavoli in un rifugio di montagna. Sotto ad uno dei tavoli occupato da alcuni turisti, un neonato sta riposando nella sua culla.]

Foreseeable: Helen knows that under the table there is a baby because she saw it; Helen just wants to wait tables. [Elena sa che sotto la tavola c'è il bambino perché lo vede, e vuole solamente servire ai tavoli.]

Or

Not foreseeable: Helen does not know that under the table there is a baby because she did not see it; Helen just wants to wait tables. [Elena non sa che sotto la tavola c'è il bambino perché non l'ha visto, e vuole solamente servire ai tavoli.]

Careful: Helen moves carefully among the tables. However, Helen steps accidentally on the cradle and hits the baby. [Elena si muove tra i tavoli prestando molta attenzione ai propri movimenti. Elena, tuttavia, pesta accidentalmente la culla e ferisce il bambino.]

Or

Careless: Helen moves carelessly among the tables. Helen steps accidentally on the cradle and hits the baby. [Elena si muove tra i tavoli prestando poca attenzione ai propri movimenti. Elena pesta accidentalmente la culla e ferisce il bambino.]

Appendix 2 - Moral Scenarios of Study 2

Scenarios were administered in Italian. Here below we provide the English translation complemented with the original version in Italian in square brackets.

Injustice Context

Background: During the summer, Jack and Thomas, who are strangers, work a few days together for a company that organizes summer camps. During the distribution of wages, the boss delivers all the money to Jack as Thomas is sick and the boss thinks that the two employees are old friends. [D'estate, Giovanni e Tommaso, pur non conoscendosi, lavorano insieme qualche giorno per una compagnia che organizza campi estivi. Al momento di ritirare la paga, poiché Tommaso è malato, il capo consegna tutti i soldi a Giovanni, pensando che i due lavoratori si conoscano da tempo.]

Neutral intention: Jack wants to keep all the money for himself, because he erroneously thinks that the wage is just for him. [Giovanni vuole prendere e tenere tutti i soldi per sé, poiché pensa che siano per lui.]

Careless: Jack takes the money absentmindedly. [Giovanni prende i soldi distrattamente.]

Or

Careful: Jack takes the money carefully. [Giovanni prende i soldi con attenzione.]

Negative outcome: Eventually, Jack keeps also Thomas' money. [Alla fine, Giovanni si tiene anche i soldi di Tommaso.]

Malevolence Context

Background: James is getting on the bus, and a very old man is also getting on with him. [Giacomo sale sull'autobus, e insieme a lui sale un uomo molto anziano.]

Neutral intention: James does not have any particular intentions; he just wants to occupy the last free seat because he is tired. [Giacomo non ha intenzioni particolari, vuole solamente sedersi nell'unico posto libero perché è stanco.]

Careless: James does not rush and sits without any concern for arriving first at the seat. [Giacomo non si precipita e si siede senza curarsi troppo di arrivare per primo al posto individuato.]

Or

Careful: James sits making sure to arrive at the seat before everyone else. [Giacomo si siede curandosi di arrivare prima di tutti al posto individuato.]

Negative outcome: James takes the seat. The old man is forced to travel standing near the entrance door. [Giacomo, sedendosi, occupa l'unico posto libero; l'uomo anziano è costretto a viaggiare in piedi.]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abelard, P. (XII century/1971). *Ethics*. Oxford: Oxford University Press.
- Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin*, 126, 556–574.

- Alicke, M. D., Buckingham, J., Zell, E., & Davis, T. (2008). Culpable control and counterfactual reasoning in the psychology of blame. *Personality and Social Psychology Bulletin*, *34*, 1371–1381.
- Alicke, M. D., Rose, D., & Bloom, D. (2011). Causation, norm violation, and culpable control. *The Journal of Philosophy*, *108*, 670–696.
- Anscombe, G. E. M. (1957). *Intention*. Oxford: Blackwell.
- Baron-Cohen, S., Tager-Flusberg, H., & Lombardo, M. (Eds.). (2013). *Understanding other minds: Perspectives from developmental social neuroscience*. Oxford: Oxford University Press.
- Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M. T., Fitzpatrick, S., Gurven, M., ... Laurence, S. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences*, *113*, 4688–4693.
- Beebe, J. R. (2013). A Knobe effect for belief ascriptions. *Review of Philosophy and Psychology*, *4*, 235–258.
- Buon, M., Dupoux, E., Jacob, P., Chaste, P., Leboyer, M., & Zalla, T. (2013). The role of causal and intentional judgments in moral reasoning in individuals with high functioning autism. *Journal of Autism and Developmental Disorders*, *43*, 458–470.
- Buon, M., Seara-Cardoso, A., & Viding, E. (2016). Why (and how) should we study the interplay between emotional arousal, theory of mind, and inhibitory control to understand moral cognition? *Psychonomic Bulletin & Review*, *23*, 1660–1680.
- Cushman, F. (2015). Deconstructing intent to reconstruct morality. *Current Opinion in Psychology*, *6*, 97–103.
- Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition*, *127*, 6–21.
- Cushman, F. A. (2008). Crime and punishment: Distinguishing the role of causal and intentional analyses in moral judgment. *Cognition*, *108*, 353–380.
- D'Arcey, E. (1963). *Human acts: An essay in their moral evaluation*. Oxford: Clarendon.
- Darley, J. M., & Shultz, T. R. (1990). Moral rules: Their content and acquisition. *Annual Review of Psychology*, *41*, 525–556.
- Dunfield, K. A., & Kuhlmeier, V. A. (2010). Intention-mediated selective helping in infancy. *Psychological Science*, *21*, 523–527.
- Enzle, M. E., & Hawkins, W. L. (1992). A priori actor negligence mediates a posteriori outcome effects on moral judgment. *Journal of Experimental Social Psychology*, *28*, 169–185.
- Finkel, N., & Groscup, J. L. (1997). When mistakes happen: Commonsense rules of culpability. *Psychology, Public Policy, and Law*, *3*, 65–125.
- Hamlin, J. K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition*, *128*, 451–474.
- Hamlin, J. K., Ullman, T., Tenenbaum, J., Goodman, N., & Baker, C. (2013). The mentalistic basis of core social cognition: Experiments with preverbal infants and a computational model. *Developmental Science*, *16*, 209–226.
- Hart, H. L. (1968). *Punishment and responsibility*. Oxford: Clarendon Press.
- Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.
- Kant, I. (1785/1959). *Foundation of the metaphysics of morals*. Indianapolis: Bobbs-Merrill.
- Karlovac, M., & Darley, J. M. (1988). Attribution of responsibility for accidents: A negligence law analogy. *Social Cognition*, *6*, 287–318.
- Kneer, M., & Machery, E. (2019). No luck for moral luck. *Cognition*, *182*, 331–348.
- Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis*, *63*, 190–193.
- Knobe, J. (2004). Intention, intentional action and moral considerations. *Analysis*, *64*, 181–187.
- Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, *108*, 754–770.
- Leslie, A. M., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: Theory of mind and moral judgment. *Psychological Science*, *5*, 421–427.
- Malle, B. F., Guglielmo, S., & Monroe, A. E. (2014). A theory of blame. *Psychological Inquiry*, *25*, 147–186.
- Malle, B. F., & Knobe, J. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, *33*, 101–121.
- Marcus Aurelius (II century AD/2006). *Meditations*. UK: Penguin Classics.
- Margoni, F., Geipel, J., Hadjichristidis, C., & Surian, L. (2018). Moral judgment in old age: Evidence for an intent-to-outcome shift. *Experimental Psychology*, *65*, 105–114.
- Margoni, F., Geipel, J., Hadjichristidis, C., & Surian, L. (2019a). The influence of agents' negligence in shaping younger and older adults' moral judgment. *Cognitive Development*, *49*, 116–126.
- Margoni, F., Guglielmetti, G., & Surian, L. (2019b). Young children with autism can generate intent-based moral judgments. *Journal of Autism and Developmental Disorders*, *49*, 5078–5085.
- Margoni, F., & Surian, L. (2016). Explaining the U-shaped development of intent-based moral judgments. *Frontiers in Psychology*, *7*, 219.
- Margoni, F., & Surian, L. (2017). Children's intention-based moral judgments of helping agents. *Cognitive Development*, *41*, 46–64.
- Margoni, F., & Surian, L. (2020). Conceptual continuity in the development of intent-based moral judgment. *Journal of Experimental Child Psychology*, *194*, 104812.
- Monroe, A. E., & Malle, B. F. (2017). Two paths to blame: Intentionality directs moral information processing along two distinct tracks. *Journal of Experimental Psychology: General*, *146*, 123–133.
- Mulvey, K. L., Gönültaş, S., & Richardson, C. B. (2020). Who is to blame? Children's and adults' moral judgments regarding victim and transgressor negligence. *Cognitive Science*, *44*, e12833.
- Nelson-le Gall, S. A. (1985). Motive–outcome matching and outcome foreseeability: Effects on attribution of intentionality and moral judgments. *Developmental Psychology*, *21*, 332–337.
- Nichols, S., & Ulatowski, J. (2007). Intuitions and individual differences: The Knobe effect revisited. *Mind & Language*, *22*, 346–365.
- Nobes, G., Panagiotaki, G., & Engelhardt, P. E. (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.
- 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*(4), 382–397.
- Pellizzoni, S., Giroto, V., & Surian, L. (2010). Belief and moral valence affect intentionality attributions: The case of side effects. *Review of Philosophy and Psychology*, *1*, 201–209.
- Pellizzoni, S., Siegal, M., & Surian, L. (2009). Foreknowledge, caring, and the side-effect effect in young children. *Developmental Psychology*, *45*, 289–295.
- Piaget, J. (1932). *The moral judgment of the child*. London: Kegan Paul.
- Proft, M., & Rakoczy, H. (2018). The ontogeny of intent-based normative judgments. *Developmental Science*, *22*, e12728.
- Reamer, F. G., & Racette, M. J. (2015). *Risk management in social work: Preventing professional malpractice, liability, and disciplinary action*. New York: Columbia University Press.
- Schleifer, M., Shultz, T. R., & Lefebvre-Pinard, M. (1983). Children's judgments of causality, responsibility and punishment in cases of harm due to omission. *British Journal of Developmental Psychology*, *1*, 87–97.
- Shen, F. X., Hoffman, M. B., Jones, O. D., & Greene, J. D. (2011). Sorting guilty minds. *New York University Law Review*, *86*, 1306.

- Shultz, T. R., Schleifer, M., & Altman, I. (1981). Judgments of causation, responsibility, and punishment in cases of harm-doing. *Canadian Journal of Behavioural Science, 13*, 238–253.
- 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.
- Shultz, T. R., Wright, K., & Schleifer, M. (1986). Assignment of moral responsibility and punishment. *Child Development, 57*, 177–184.
- Siegal, M., & Peterson, C. C. (1998). Preschoolers' understanding of lies and innocent and negligent mistakes. *Developmental Psychology, 34*, 332–342.
- Strid, K., & Meristo, M. (2020). Infants consider the distributor's intentions in resource allocation. *Frontiers in Psychology, 11*, 596213.
- Uttich, K., & Lombrozo, T. (2010). Norms inform mental state ascriptions: A rational explanation for the side-effect effect. *Cognition, 116*, 87–100.
- Weiner, B. (1995). *Judgments of responsibility: A foundation for a theory of social conduct*. New York: Guilford.
- Williams, G. L. (1953). *Criminal law: The general part*. London: Stevens & Sons.
- Woo, B. M., Steckler, C. M., Le, D. T., & Hamlin, J. K. (2017). Social evaluation of intentional, truly accidental, and negligently accidental helpers and harmers by 10-month-old infants. *Cognition, 168*, 154–163.
- Young, L., & Saxe, R. (2009a). An fMRI investigation of spontaneous mental state inference for moral judgment. *Journal of Cognitive Neuroscience, 21*, 1396–1405.
- Young, L., & Saxe, R. (2009b). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia, 47*, 2065–2072.
- Zalla, T., & Leboyer, M. (2011). Judgment of intentionality and moral evaluation in individuals with high functioning autism. *Review of Philosophy and Psychology, 2*, 681–698.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.