The influence of agents' negligence in shaping younger and older adults' moral judgment

Francesco Margoni

Janet Geipel

Constantinos Hadjichristidis

Luca Surian

The article is accepted for publication in *Cognitive Development* – 19 December 2018

Author Note

Francesco Margoni and Luca Surian, Department of Psychology and Cognitive Sciences, University of Trento. Janet Geipel, Department of Psychology, The University of Chicago. Constantinos Hadjichristidis, Department of Economics and Management, University of Trento, and Centre for Decision Research, Leeds University Business School, University of Leeds.

Correspondence concerning this article should be addressed to Francesco Margoni, University of Trento, Department of Psychology and Cognitive Sciences, E-mail: francesco.margoni@unitn.it, or to Janet Geipel, Department of Psychology, The University of Chicago, E-mail: jgeipel@uchicago.edu

Abstract

This study examined age-related differences in the use of negligence information in moral judgment. A group of younger adults (18–36 years) and a group of older adults (75–98 years) were presented with a series of scenarios illustrating cases where an agent unintentionally causes harm. The scenarios also specified whether or not the agent acted with negligence. Participants were asked to rate how morally wrong was the agent's action. We found that older participants condemned the agents of accidental harms regardless of whether they acted with negligence, whereas younger participants condemned only the agents that acted with negligence. Subsequently, participants were presented with an accidental harm scenario in which negligence information was omitted, and were asked to morally evaluate the agent's action and to rate the extent to which the agent could be accused of negligence. Compared to younger adults, older adults condemned the agent's action more severely and rated the agent as more negligent. These results suggest that aging is associated with an increased tendency to assume that accidental harmdoers are negligent. This bias may help explain the intent-to-outcome shift occurring in old adults' moral judgment.

Keywords: moral judgment, negligence, intention, older adults, cognition

The influence of agents' negligence in shaping younger and older adults' moral judgment

People morally evaluate human actions by attending to both their external outcomes and the underlying intentions. However, the relative weight people assign to these types of information changes throughout the lifespan. Classical works in the developmental literature suggest that young children attend more to outcomes than to intentions whereas older children show the opposite bias (Piaget, 1932). More recently, studies have shown that older adults manifest a reversal of this developmental shift (Margoni, Geipel, Hadjichristidis, & Surian, 2018; Moran, Jolly, & Mitchell, 2012). For example, when presented with a case of accidental harm (unintentionally harming someone), older adults tended to morally condemn the agent that was causally responsible for the harm, suggesting that their attention was focused more on the presence of a negative outcome than on the absence of a negative intention. By contrast, younger adults tended to focus more on intention, exculpating agents who accidentally caused harm (Cushman, 2008, 2013). Thus, while during childhood we observe an *outcome-to-intent shift*, in old age we observe an *intent-to-outcome shift*.

Apart from intentions and outcomes, mature moral reasoners also pay attention to negligence. It is often said that an agent was negligent if he or she acted without due care (Alicke, 1992, 2000; Enzle & Hawking, 1992; Monroe & Malle, 2017; Schleifer, Shultz, & Lefebvre-Pinard, 1983; Shultz & Wright, 1985; Shultz, Wright, & Schleifer, 1986; Siegal & Peterson, 1998). If someone unintentionally caused a bad outcome but acted with negligence, adults tend to morally condemn the agent despite the absence of a bad intention. The aims of the present study were: (a) to investigate age-related differences in the tendency to take negligence into account by comparing average moral wrongness and punishment judgments of a group of younger adults to those of a group of older adults, and (b) to assess whether these differences help explain age-related differences in the tendency to take intention and outcome into account.

The *outcome-to-intent shift* occurs during the later preschool years: the majority of studies which used tasks that require a verbal response indicate that it is only by the age of five years that most children display the ability to *express* intent-based moral judgments (e.g., Cushman, Sheketoff, Wharton, & Carey, 2013; Margoni & Surian, 2017; Nobes, Panagiotaki, & Bartholomew, 2016; though see Nobes, Panagiotaki, & Engelhardt, 2017). However, studies using non-verbal response measures (see Margoni & Surian, 2018) have revealed that even infants show sensitivity to intention (e.g., Hamlin, 2013; Woo, Steckler, Le, & Hamlin, 2017). How can these results be reconciled? According to a recent account, the *expression view*, these divergent results are due to differences in the processing demands associated with verbal and non-verbal tasks. Namely, verbal tasks are more taxing on young preschoolers' limited cognitive capacity than non-verbal tasks (Margoni & Surian, 2016).

More specifically, the expression view holds that the apparent outcome-to-intent shift documented on verbal-response tasks is due to developmental changes in theory of mind or executive functioning skills (see also Buon, Seara-Cardoso, & Viding, 2016). Thus, changes occurring outside the moral domain both early (Garon, Bryson, & Smith, 2008; Slaughter, 2015) and later in life (Henry, Phillips, Ruffman, & Bailey, 2013; Reuter-Lorenz & Sylvester, 2005; Salthouse, 2004), for example changes in executive functioning skills, may help explain the developmental trajectory of the use of intention and outcome information throughout the lifespan (Margoni et al., 2018; see also Chen & Blanchard-Fields, 2000; Ligneaur-Herve & Mullet, 2005; Pratt, Diessner, Pratt, Hunsberger, & Pancer, 1996). In particular, both younger preschoolers and older adults may find it difficult to suppress cues concerning action outcomes, while older children and younger adults may possess sufficient skills to inhibit the prepotent outcome-based response in favor of an intent-based response (Buon, Jacob, Loissel, & Dupoux, 2013).

Another factor that has been considered useful in accounting for the outcome-to-intent shift is the child's attribution of negligence to agents (Nobes et al., 2017). Nobes, Panagiotaki and Pawson (2009) suggested that younger preschoolers condemn accidental harms not simply because they focus on the negative outcomes, but because they over-attribute negligence to the accidental transgressor. To the extent that this account is right, the outcome-to-intent shift may be a misnomer, as it may simply reflect a developmental change in the inferences young children make about negligence. Specifically, as they develop, young children move from an over-attribution of negligence to a more proper utilization of negligence information (notice that the developmental trajectory for intentions follows the opposite direction: from an underutilization towards a more proper utilization). Nobes and colleagues (2009) suggest that younger children's over-attribution of negligence might stem from their belief that all negative outcomes are foreseeable and therefore avoidable. Accidental transgressors are blameworthy because they should have foreseen the negative outcomes of their actions.

We hypothesize that older adults might over-attribute negligence for similar reasons: they might also perceive negative outcomes as foreseeable. In support of this hypothesis, studies have shown that in comparison to younger adults, both younger children and older adults are more susceptible to hindsight bias, which refers to the tendency to see something as inevitable once it has occurred (Bernstein, Erdfelder, Meltzoff, Peria, & Loftus, 2011). The negligence view can be integrated with the expression account detailed above. Considering alternative ways in which the future might have unfolded necessitates executive functioning skills. Thus, one may posit that attributing causal responsibility based on outcomes is automatic, whereas properly integrating intention and negligence information is cognitively taxing (Buon et al., 2013). In relation to accidental harms, the attribution of causal responsibility to the transgressor may be the default,

whereas the proper consideration of negligence and intentionality information may require more mental effort. Due to these default inferences, younger preschoolers and older adults may be more condemning of accidental transgressors' actions.

Present Study

In the present study we employed two tasks to investigate age-related differences in the extent to which negligence information (either stated explicitly or inferred) is integrated in moral judgment. In Task 1, participants evaluated the moral wrongness and punishability of cases of accidental harm in which information about the agent's negligence, or lack of it, was explicitly stated. We predicted that younger adults would tend to condemn accidental transgressors who acted with negligence and to exculpate those who acted without negligence. Also, we expected that older adults' judgments would be less affected by the explicit information about the absence of agents' negligence, and, because they tend to focus more on outcome information than on mental state information, they would display a tendency to condemn also transgressors that are described as non-negligent.

In Task 2, participants evaluated the moral wrongness and punishability of cases of accidental harm in which no information was provided about whether the agent acted with or without negligence. The aim of Task 2 was to investigate younger and older adults' tendency to attribute negligence to accidental harmdoers. We predicted that older adults, as compared to younger adults, would show a higher tendency to spontaneously infer negligence in agents that caused accidental harms in the absence of explicit information about negligence. The basis of our prediction is that inferring negligence from bad outcomes is the default. Moreover, to gain a better understanding of the role of executive functioning skills on older adults' moral and

punishment judgments, we assessed participants' working memory skills and predicted that they would mediate the effect of age on moral judgment.

In both tasks, we expected to find an age group effect on both moral wrongness and punishment judgments. However, it is noteworthy to mention that moral wrongness and punishment judgments tap into different underlying constructs: blame and punishment, respectively. Consistent with this distinction, empirical evidence shows that accidental harms are judged more leniently in terms of moral wrongness than in terms of punishability (Cushman, 2008). This is presumably because moral wrongness judgments are influenced mostly by intentions, whereas punishment judgments are influenced by both intentions *and* outcomes. If this is the case, then based on our hypothesis, age group should have a bigger effect on moral wrongness than on punishment judgments.

Methods

Participants

The sample size was determined by an a-priori power analysis using the Shiny webb app (Anderson, Kelley, & Maxwell, 2017) for a mixed ANOVA using an uncertainty and publication bias correction. We used the following estimates: $F = 42.02$ with a total sample size of 58 (based on Margoni et al., 2018), with two levels for the between-subject factor (old, young), two levels for the within-subject factor (with negligence, without negligence), alpha set at .05, a power of .95, and a desired level of assurance of .95. This analysis indicated a minimum sample size of 40 participants per age group. No interim or stopping rules were applied. The study protocol for Task 1 was pre-registered (http://aspredicted.org/blind.php?x=m3jv5e), but that of Task 2 was not. We decided to include Task 2 later, after we had pre-registered Task 1. Its purpose was to test whether older adults are more prone than younger adults to attribute negligence to an agent that accidentally caused some harm.

We recruited 82 participants: 41 older adults (34 female, M_{Age} = 87;0 years, SD_{Age} = 6;4, age range: 75–98 years), and 41 younger adults (26 female, M_{Age} = 24;6 years, SD_{Age} = 4;11, age range: 18–36 years). The older participants were recruited through local nursing homes, while the younger participants through flyers posted at the campus of the University of Trento. The research protocol of the study was approved by the Ethics Committee of the University of Trento ("The moral judgment in old age", protocol number 2017-015).

Materials and Procedure

Task 1 – Attending to negligence. We constructed four scenarios (by adapting the four harm scenarios used in Margoni et al., 2018; see Supplementary Materials for the complete battery). Each scenario had two experimental versions and two control versions. The two experimental versions involved harmful consequences brought about by actions that were motivated by neutral intentions. In one version the agent acted without due care (*negligence version*; e.g., Chloe sold a sick dog which was infected with rabies because she did not check the dog carefully), whereas in the other the agent acted with due care (*no-negligence version*; e.g., Chloe sold a sick dog which was infected with rabies, as a careful assessment of the dog made her believe that the dog was healthy). The two control versions were *all-neutral* or *all-bad* cases. The all-neutral cases involved an action motivated by a neutral intention that resulted in a neutral outcome (e.g., Chloe intended to sell a healthy dog and did so, as the dog was healthy), whereas the all-bad case involved an action motivated by a bad intention that resulted in a bad outcome (e.g., Chloe intended to sell a sick dog which was infected with rabies, and purposely did so). Each participant judged a different version of each of the four scenarios. Across participants, we rotated the version selection following a Latin square design.

Following each scenario, participants were asked two questions:

Moral wrongness question: "How morally wrong was [the agent's] action?";

Punishment question: "How much do you think that [the agent] should be punished?".

For each question, participants were asked to respond on an 11-point scale, ranging from 0 to 10 (0 = *Not at all*, 5 = *Somewhat*, 10 = *Very much*). The order of the test questions (moral wrongness question first vs. punishment question first) was counterbalanced across participants.

Task 2 – Inferring negligence. Task 2 was not pre-registered as we decided to include it after we had pre-registered Task 1. Its aim was twofold: (1) to replicate the main effect of Margoni et al. (2018) which showed that older adults, compared to younger adults, are more prone to morally condemn the agents of accidental harms, and (2) to examine whether this tendency is in part due to the fact that older adults are more prone to attribute negligence to agents that brought about accidental harms. A difference between Task 1 and Task 2 is that whereas in the accidental harm scenarios of Task 1 *explicit* information about negligence was provided, in the scenarios of Task 2 such information was always omitted. In relation to Task 1, Task 2 served the complementary purpose of investigating age-related differences in *spontaneous* inferences about negligence.

We selected the *accidental harm* versions of the four harm scenarios used in Margoni et al. (2018). These stories do not specify information about negligence. Following the story, which was presented immediately after the last story of Task 1, participants were asked three questions:

Moral wrongness question: "How morally wrong was [the agent's] action?";

Punishment question: "How much do you think that [the agent] should be punished?";

Negligence attribution question: "To what extent do you think that [the agent] could be accused of negligence, that is, that [she/he] was aware of the possible risks and acted without the necessary caution?".

For each question, participants were asked to respond on an 11-point scale, ranging from 0 to 10 (0 = *Not at all*, 5 = *Somewhat*, 10 = *Very much*). Finally, participants were instructed "Now, we ask you to briefly explain why according to you the action of [the agent] is not at all, somewhat, or very wrong and why you think that it is not at all, somewhat, or very punishable." We assigned value 1 if the participant mentioned negligence in his or her response, and 0 if the participant did not mention negligence. Each participant was presented with just one of the four stories (from the scenario that the participant read in Task 1 in the all-neutral version), which were counterbalanced across the age groups. The order of the three questions following the story (moral wrongness, punishment and negligence questions) was counterbalanced across participants.

Working memory skills. Following the moral judgment tasks, participants were asked to complete a listening span test (Pazzaglia, Palladino, & De Beni, 2000), which is the Italian version of the Reading Span Test (Daneman & Carpener, 1980). This test measures individual differences in listening comprehension, which may reflect differences in working memory skills. Participants were asked to say whether some sentences are true or false, directly after the experimenter finished the sentence. After an increasing number of sentences (it may be two, three, four, five or six), participants were asked to recall the last word of each sentence. The experimenter took note of the number of correct words recalled by each participant.

Results

Preliminary Analyses

Screening for dementia. The Mini-Mental Status Examination was employed as a screening tool for dementia (Folstein, Folstein, & McHugh, 1975). None of the older participants had severe impairment/dementia (scores between 0 and 9), moderate impairment (10-16) or mild

impairment (17-19). Twelve participants had scores that revealed a suspected impairment (20-24), while the rest had no impairment (24-30).

Age, education, and gender differences. There was a significant difference between the age groups in terms of gender composition, χ^2 (1, N = 82) = 3.98, p = .046, φ = .22. There was also a significant difference between the age groups in terms of years of education, t(66.94) = 11.55, p < .001, d = 2.55. On average, older participants indicated fewer years of school education (M = 7.44 years, CI [6.41, 8.47]) than younger participants (M = 14.39 years, CI [13.75, 15.03]). To examine whether gender and education influence moral wrongness judgment, we conducted two linear regression analyses, one for each age group. We found that gender (younger participants: β = –.17, t(39) = –1.07, p = .292; older participants: β = –.14, t(39) = –0.86, p = .397) and education (younger participants: β = .05, t(39) = 0.34, p = .739; older participants: β = –.22, t(39) = –1.36, p = .181) did not significantly predict moral wrongness judgment in either age group. Next, to examine whether gender and education influence punishment judgment, we conducted two linear regression analyses, one for each age group. We found that gender (younger participants: β = –.08, t(39) = –0.52, p = .608; older participants: β = –.08, t(38) = –0.50, p = .621) and education (younger participants: β = .29, t(39) = 1.86, p = .071; older participants: β = –.09, t(38) = –0.56, p = .577) did not significantly predict punishment judgment in either age group. Thus, we omitted gender and education from subsequent analyses.

Order of question effects. Preliminary analyses revealed that presentation order did not interact with neither age group or negligence level, Fs < 1.25, ps > .25. In light of this finding, and in order to simplify the analyses, we collapsed the data over presentation order.

Task 1

We first report the results for moral wrongness judgments and punishment judgments separately. However, we also report a joint analysis to detect whether: (a) accidental harms are judged more leniently in terms of moral wrongness than in terms of punishability and (b) whether the age group difference is higher for moral wrongness judgments. Note that we preregistered only the main analyses, that is those concerning the experimental items (accidental harm scenarios with or without negligence) for which we expected an age difference. However, our aim from the very beginning was also to test for age differences with the control scenarios for which we expected no age differences.

Moral wrongness judgment. We first focused on the two accidental harm scenarios. We predicted that older participants, unlike younger participants, would condemn both negligent and non-negligent transgressors. The main results from Task 1 are illustrated in Figure 1. We submitted the moral wrongness judgments to a 2 (Age group: old vs. young) × 2 (Negligence level: present vs. absent) mixed-factor ANOVA, with repeated measures on negligence level. The analysis revealed a significant main effect of negligence level, $F(1, 80) = 16.72$, $p < .001$, $\eta_p^2 = .17$, $f = 0.46$. Accidental harms that resulted from negligence were judged as more morally wrong ($M = 5.88$, 95% CI [5.02, 6.74]) than accidental harms that did not result from negligence ($M = 3.96$, CI [3.17, 4.76]). Critically, this effect was qualified by an Age group × Negligence level interaction, $F(1, 80) = 4.45$, $p = .038$, $\eta_p^2 = .05$, $f = 0.24$. Pairwise tests, adjusted for multiple comparisons, showed that older participants were less influenced by negligence level ($M_{Present} = 7.05$, CI [5.83, 8.27]; $M_{Absent} = 6.12$, CI [5.00, 7.25]), $F(1, 80) = 1.96$, $p = .166$, $\eta_p^2 = .02$, $f = 0.16$, than younger participants ($M_{Present} = 4.71$, CI [3.49, 5.93]; $M_{Absent} = 1.81$, CI [0.68, 2.93]), $F(1, 80) = 19.21$, $p < .001$, $\eta_p^2 = .19$, $f = 0.48$. There was also a significant main effect of age group, $F(1, 80) = 23.31$, $p < .001$, $\eta_p^2 = .23$, $f = 0.55$. Overall, older participants gave higher

moral wrongness ratings (M = 6.59, CI [5.62, 7.56]) than younger participants (M = 3.26, CI [2.29, 4.22]).

Next, we focused on the control versions of the scenarios. We submitted the moral wrongness judgments to a 2 (Age group: old vs. young) × 2 (Scenario: all-bad vs. all-neutral) mixed-factor ANOVA, with repeated measures on scenario. The only significant effect was a main effect of scenario, $F(1, 80) = 580.77$, $p < .001$, $\eta_p^2 = .88$, $f = 2.70$. Overall, the all-bad version was judged as more morally wrong (M = 8.37, CI [7.75, 8.98]) than the all-neutral version (M = 0.68, CI [0.34, 1.03]). These results suggest that the participants paid attention, and help rule out alternative interpretations for the finding concerning accidental harms such as that older adults make a different use of the rating scale (use higher ratings). The analysis revealed no main effect of age group, $F(1, 80) = 0.58$, $p = .449$, $\eta_p^2 < .01$, $f < 0.10$, nor was there an Age group × Scenario interaction, $F(1, 80) = 0.84$, $p = .361$, $\eta_p^2 = .01$, $f = 0.10$.

Punishment judgment. We first submitted the punishment ratings of the accidental harm scenarios to a 2 (Age group: old vs. young) × 2 (Negligence level: present vs. absent) mixed-factor ANOVA, with repeated measures on negligence level. The analysis revealed a significant main effect of negligence level, $F(1, 79) = 15.78$, $p < .001$, $\eta_p^2 = .17$, $f = 0.45$, which was qualified by a marginally significant Age group × Negligence level interaction, $F(1, 79) = 3.71$, $p = .058$, $\eta_p^2 = .05$, $f = 0.22$. As was the case with moral wrongness judgments, pairwise comparisons showed that older participants were less influenced by negligence level ($M_{Present}$ = 7.38, CI [6.30, 8.44]; M_{Absent} = 6.30, CI [5.11, 7.49]), $F(1, 79) = 2.07$, $p = .154$, $\eta_p^2 = .03$, $f = 0.16$, than were younger participants ($M_{Present}$ = 6.10, CI [5.04, 7.16]; M_{Absent} = 3.00, CI [1.83, 4.17]), $F(1, 79) = 17.61$, $p < .001$, $\eta_p^2 = .18$, $f = 0.47$. There was also a main effect of age group,

$F(1, 79) = 14.48$, $p < .001$, $\eta_p^2 = .16$, $f = 0.44$. Overall, older participants gave higher punishment ratings ($M = 6.84$, CI [5.99, 7.69]) than younger participants ($M = 4.55$, CI [3.71, 5.39]).

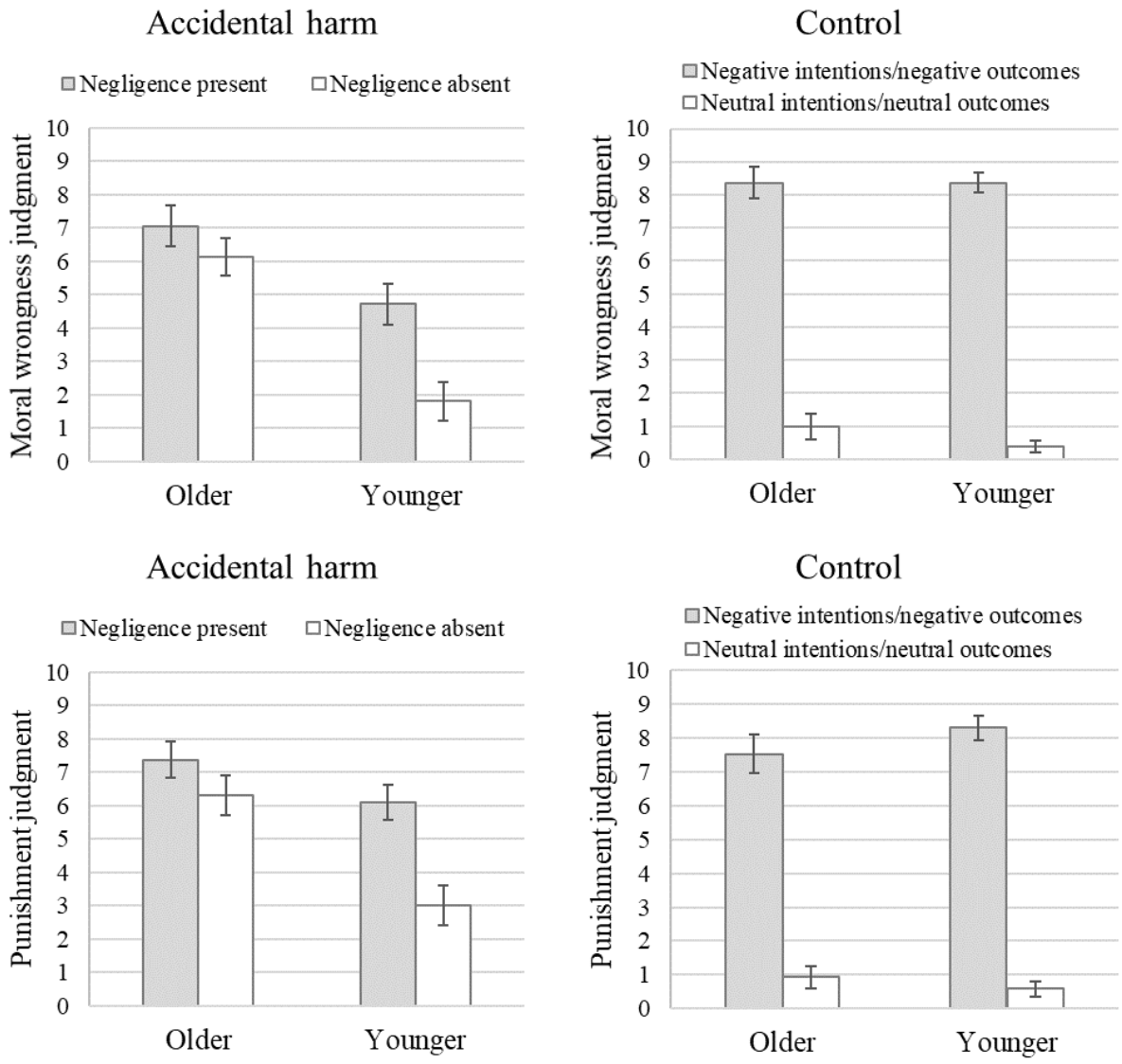

Figure 1. Average moral wrongness ratings (top panel) and average punishment ratings (bottom panel) by age group and negligence level for accidental harm scenarios (left) or scenario type for control scenarios (right).

Next, we submitted the punishment ratings of the two control scenarios to a 2 (Age group: old vs. young) × 2 (Scenario: all-neutral vs. all-bad) mixed-factor ANOVA, with repeated

measures on scenario. There was only a main effect of scenario, $F(1, 80) = 401.65$, $p < .001$, $\eta_p^2 = .83$, $f = 2.20$. Overall, the all-bad control scenario version received higher punishment ratings ($M = 7.93$, CI [7.24, 8.62]) than the all-neutral scenario version ($M = 0.76$, CI [0.35, 1.16]). This result suggests that participants were attentive. The analysis revealed no main effect of age group, $F(1, 80) = 0.25$, $p = .621$, $\eta_p^2 < .01$, $f = 0.05$, nor was there an Age group × Scenario interaction, $F(1, 80) = 2.46$, $p = .121$, $\eta_p^2 = .03$, $f = 0.18$.

Joint analysis of moral wrongness judgments and punishment judgments. The purpose of this analysis was: (a) to test whether accidental harms receive more severe punishment than moral wrongness judgments (Cushman, 2008), and (b), in relation to this, whether the age group effect is less pronounced for punishment than for moral wrongness judgments.

We submitted the moral wrongness and punishment judgments to a 2 (Age group: old vs. young) × 2 (Negligence level: present vs. absent) × 2 (Judgment type: moral wrongness vs. punishment) mixed-factor ANOVA, with repeated measures on negligence level and judgment type. In relation to (a), we found a significant main effect of judgment type, $F(1, 79) = 15.97$, $p < .001$, $\eta_p^2 = .17$, $f = 0.45$, such that accidental harms received more severe punishment ratings ($M_{Punishment} = 5.69$, CI [5.10, 6.29]) than moral wrongness ratings ($M_{Wrongness} = 4.88$, CI [4.19, 5.57]). In relation to (b), the effect of judgment type was qualified by a Age group × Judgment type interaction, $F(1, 79) = 5.48$, $p = .022$, $\eta_p^2 = .07$, $f = 0.26$. The age group effect was significant in both cases but less pronounced in punishment judgments ($M_{Young} = 4.55$, CI [3.71, 5.39] vs. $M_{Old} = 6.84$, CI [5.99, 7.69], $M_{Diff} = 2.29$), $F(1, 79) = 14.48$, $p < .001$, $\eta_p^2 = .16$, $f = 0.42$, than in moral wrongness judgments ($M_{Young} = 3.26$, CI [2.29, 4.23] vs. $M_{Old} = 6.50$, CI [5.52, 7.48], $M_{Diff} = 3.24$), $F(1, 79) = 21.91$, $p < .001$, $\eta_p^2 = .21$, $f = 0.52$.

The results of this analysis replicated the results from the separate analyses reported above. There was a significant main effect of age group, $F(1, 79) = 20.17$, $p < .001$, $\eta_p^2 = .20$, $f = 0.50$, a significant main effect of negligence level, $F(1, 79) = 18.38$, $p < .001$, $\eta_p^2 = .19$, $f = 0.48$, and, crucially, the predicted Age group × Negligence level interaction, $F(1, 79) = 4.51$, $p = .037$, $\eta_p^2 = .05$, $f = 0.24$. There was no Age group × Negligence level × Judgment type interaction, $F(1, 79) = 0.01$, $p = .921$, $\eta_p^2 < .01$, $f < 0.10$.

Correlations between age group, moral wrongness judgment, punishment judgment, and working memory. The following analyses pertain to accidental harm scenarios. Table 1 shows zero-order correlations between age group, moral wrongness judgments for scenarios with negligence, scenarios without negligence, and their difference ($D_{Wrongness}$ = average wrongness rating for scenarios with negligence *minus* average wrongness rating for scenarios without negligence), punishment judgments for scenarios with negligence, scenarios without negligence, and their difference ($D_{Punishment}$), and working memory skills (we used the standardized z scores). Age was entered as a binary variable (0 = younger participants, 1 = older participants). Age group was negatively correlated both with $D_{Wrongness}$ and with working memory skills, while $D_{Wrongness}$ was marginally correlated with working memory skills. Age group was not significantly correlated with $D_{Punishment}$, while $D_{Punishment}$ was positively correlated with working memory skills.

Table 1.

Correlations Between Age Group, Moral Wrongness Judgment, Punishment Judgment, and Working Memory Skills.

	1	**2a**	**2b**	**3**	**4a**	**4b**	**5**	**6**
1. Age group	--							
2. *Moral wrongness judgment* (a) Accidental harm with negligence	**.29****	--						
(b) Accidental harm without negligence	**.52****	**.45****	--					
3. Difference between 2a and 2b	**–.23***	**.50****	**–.55****	--				
4. *Punishment judgment* (a) Accidental harm with negligence	.19†	**.72****	**.22***	**.46****	--			
(b) Accidental harm without negligence	**.41****	**.36****	**.89****	**–.51****	.20†	--		
5. Difference between 4a and 4b	–.21†	.21†	**–.60****	**.77****	**.55****	**–.71****	--	
6. Working memory skills	**–.71****	**–.34****	**–.49****	.15	–.16	**–.40****	**.22***	--

Note. ***p* < .01, **p* < .05, †*p* < .10.

Relationship between age group, net negligence score, and working memory. We next assessed whether age group differences in working memory skills contribute to age group differences in moral wrongness judgments. We used 1,000 bootstrapping resamples (Preacher & Hayes, 2008). As our outcome variable, we used $D_{\text{Wrongness}}$, age group was entered as the independent variable (0 = younger participants, 1 = older participants), and working memory skills as the mediator.

The relationship between age group and the net negligence score was not mediated by working memory skills. As Figure 2 illustrates, the unstandardized regression coefficient between age group and working memory skills was significant, $b = –2.33$, $p < .001$, 95% CI [–2.85, –1.82], however, the unstandardized regression coefficient between working memory skills and moral wrongness judgment was not, $b = –0.06$, $p = .881$, 95% CI [–0.88, 0.76]. The partially standardized indirect effect was $\beta = .03$, 95% [–0.60, 0.52]. The bootstrapped unstandardized indirect effect was $b = 0.14$, 95% [–2.63, 2.21], thus, the indirect effect was not statistically significant.

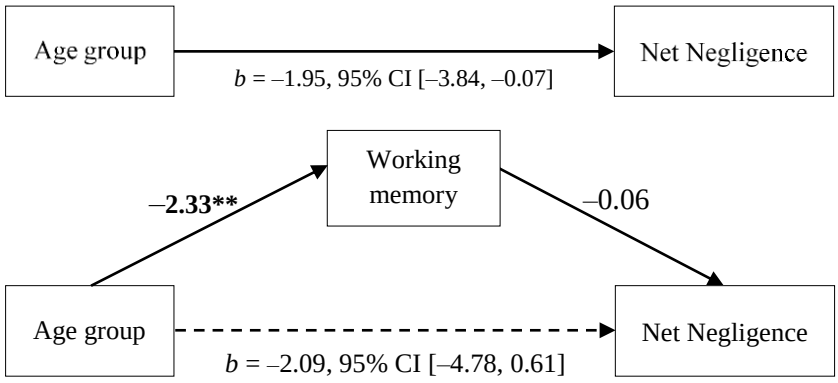

Figure 2. Unstandardized regression coefficients and bootstrap confidence intervals for the relationship between age group and net negligence as mediated by working memory skills.

Task 2

Moral wrongness judgment. We submitted the moral wrongness ratings to a simple one-way ANOVA. The analysis revealed significant age group differences, $F(1, 78) = 38.08$, $p < .001$, $\eta^2 = .33$, $f = 0.70$. Older participants assigned higher moral wrongness ratings to the accidental harm scenario ($M = 7.28$, CI [6.18, 8.37]) than did younger participants ($M = 2.48$, CI [1.38, 3.57]).

Punishment judgment. An one-way ANOVA revealed significant age group differences, $F(1, 78) = 23.80$, $p < .001$, $\eta^2 = .23$, $f = 0.55$. Older participants assigned higher punishment ratings ($M = 7.48$, CI [6.40, 8.55]) than did younger participants ($M = 3.75$, CI [2.68, 4.83]).

Perceived negligence judgment. An one-way ANOVA revealed significant age group differences, *Welch's* $F(1, 71.75) = 37.56$, $p < .001$, $\eta^2 = .33$, $f = 0.69$ (we report the Welch adjusted F-ratio because the assumption of homogeneity of variance was not met). In the absence

of negligence information, older participants rated the agent as more negligent (M = 8.18, CI [7.18, 9.17]) than younger participants (M = 3.85, CI [2.86, 4.84]).

Joint analysis of moral wrongness judgments and punishment judgments. We submitted the moral wrongness and punishment judgments to a 2 (Age group: old vs. young) × 2 (Judgment type: moral wrongness vs. punishment) mixed-factor ANOVA, with repeated measures on judgment type. There was a significant main effect of age group, $F(1, 78) = 32.97$, $p < .001$, $\eta_p^2 = .30$, $f = 0.65$, such as that older participants gave harsher judgments (M_{Old} = 7.38, CI [6.33, 8.42]) than younger participants (M_{Young} = 3.11, CI [2.07, 4.16]). There was also a significant main effect of judgment type, $F(1, 78) = 12.65$, $p = .001$, $\eta_p^2 = .14$, $f = 0.40$, such that participants gave harsher punishment ratings ($M_{Punishment}$ = 5.61, CI [4.85, 6.37]) than moral wrongness ratings ($M_{Wrongness}$ = 4.88, CI [4.10, 5.65]). This main effect was qualified by a Age group × Judgment type interaction, $F(1, 78) = 6.72$, $p = .011$, $\eta_p^2 = .08$, $f = 0.29$. The age group effect was less pronounced for punishment judgments (M_{Young} = 3.75, CI [2.68, 4.83], M_{Old} = 7.48, CI [6.40, 8.55], M_{Diff} = 3.73) as compared to moral wrongness judgments (M_{Young}= 2.48, 95% CI [1.38, 3.57], M_{Old} = 7.28, CI [6.18, 8.37], M_{Diff} = 4.80).

Explanation for moral wrongness and punishment judgments. We analyzed participants' explanations behind their moral wrongness and punishment judgments with a chi-square test (i.e., 1 = participant mentioned negligence, 0 = participant did not mention negligence). The interrater reliability of two raters was high, .91 (disagreement were resolved by a brief discussion). While 78.4% of older adults indicated negligence as a reason, only 52.6% of younger adults did so, $\chi^2 (1, N = 75) = 5.49$, $p = .019$, $\varphi = 0.27$.

Correlation between age group, moral wrongness judgment, punishment judgment, perceived negligence and working memory. Table 2 shows correlations between age group,

moral wrongness judgments, punishment judgments, perceived negligence and working memory. Age was entered as a binary variable as before. Moral wrongness judgment of the accidental harm scenario was positively correlated with age group and perceived negligence, but negatively correlated with working memory skills. Punishment judgment of the accidental harm scenario was positively correlated with age group and perceived negligence, but negatively correlated with working memory skills. Age group was negatively correlated with working memory skills which was negatively correlated with perceived negligence.

Table 2.

Correlations Between Age Group, Moral Wrongness Judgment, Punishment Judgment, Perceived Negligence and Working Memory Skills.

	1	**2**	**3**	**4**	**5**
1. Age group	**--**				
2. Accidental harm moral wrongness	**.57****	**--**			
3. Accidental harm punishment	**48****	**.89****	**--**		
4. Accidental harm perceived negligence	**.57****	**.83****	**.81****	**--**	
5. Working memory skills	**–.71****	**–.51****	**–.46****	**–.48****	**--**

Note. ***p* < .01, **p* < .05.

Relationship between age group, moral wrongness judgment, perceived negligence and working memory. The analyses above are consistent with the hypothesis that older adults, compared to younger adults, judged accidental harms more harshly, because they were more likely to ascribe negligence to the agent. We examined this hypothesis by running a mediation analysis where the outcome variable was moral wrongness judgment, the independent variable was age group, and the mediators were working memory skills and perceived negligence.The relationship between age group (0 = younger adults, 1 = older adults) and moral wrongness judgment was mediated by perceived negligence and working memory skills. As Figure 3

illustrates, the unstandardized regression coefficient between age group and perceived negligence was significant, $b = 4.28$, $p < .001$, 95% CI [2.86, 5.70], as was the unstandardized regression coefficient between age group and working memory skills, $b = –2.35$, $p < .001$, 95% CI [–2.88, –1.82]. The unstandardized regression coefficient between perceived negligence and moral wrongness judgment was also significant, $b = 0.80$, $p < .001$, 95% CI [0.63, 0.97], however, the unstandardized regression coefficient between working memory skills and moral wrongness judgment was not significant, $b = –0.24$, $p = .303$, 95% CI [–0.69, 0.22]. The partially standardized indirect effect was β = .95, 95% [0.55, 1.43]. The bootstrapped unstandardized indirect effect was $b = 3.99$, 95% [2.25, 6.17], thus, the indirect effect was statistically significant. Investigating the specific indirect effects, perceived negligence significantly explained some of the variance of the effect of age group on moral wrongness judgment, $b = 3.43$, 95% CI [2.22, 5.14] (β = .82, 95% CI [0.54, 1.16]), while working memory skills did not, $b = 0.55$, 95% CI [–0.70, 1.80] (β = .13, 95% CI [–0.16, 0.44]).

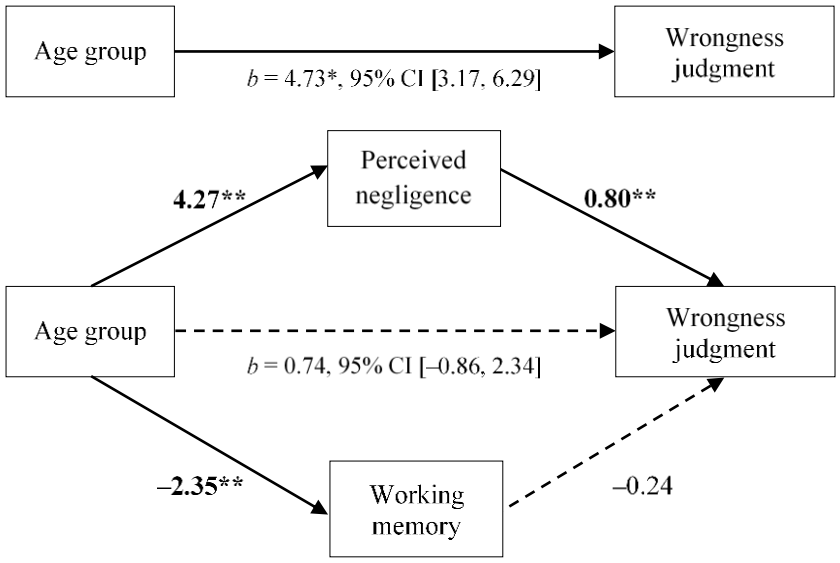

Figure 3. Unstandardized regression coefficients and bootstrap confidence intervals for the association between age group and moral wrongness judgment for accidental harm scenarios as mediated by perceived negligence and working memory skills.

Discussion

We examined whether normal aging influences people's reliance on negligence information in their moral wrongness and punishment judgments. In Task 1, we used accidental harm scenarios that explicitly stated whether or not the agent acted with negligence. We found that while younger adults condemned more severely negligent than non-negligent agents, older participants condemned equally severely both negligent and non-negligent agents. Importantly, we found no age-related differences in control scenarios in which both intentions and outcomes were either neutral or bad. In Task 2, we used an accidental harm scenario that omitted negligence information. We found that older adults condemned the accidental trangressors more than did younger adults, and were more likely to attribute negligence to the actions. Further analyses showed that perceived negligence, but not working memory skills, mediated the relationship between age group and moral wrongness judgment.

The present results on age-related differences in the use of negligence information in moral judgment help explain recent findings suggesting the occurrence of an intent-to-outcome developmental shift later in life (Margoni et al., 2018; Moran et al., 2012). Taken together with previous findings, the results suggest that older adults rely more on outcomes than on intentions in their moral evaluations, and are more likely to attribute negligence to accidental transgressors. The older adults' greater tendency to spontaneously attribute negligence to accidental transgressors found in Task 2 may have played a role in their increased condemnation of

accidental trangressors in Task 1. That is, it could be that in Task 1 older adults did not rely on the explicit information about the absence of agents' negligence and may have instead relied on their own attribution of negligence to accidental transgressors (for similar findings with younger preschoolers, see Nobes et al., 2009). In line with an account that places emphasis on inhibitory capacities, older adults may have been more prone to condemn accidental harms also because they were incapable of inhibiting their own inferences about negligence. The current findings highlight the role of negligence in the processing of moral scenarios: It is not simply that older adults selectively focus on negative outcomes; critically, they also assume that the agents behind these outcomes are negligent.

Moreover, we reported that differences between younger and older adults in judging accidental harms by relying on outcomes independently on how the agent was described by the experimenter (negligent or non-negligent) were more pronounced for moral wrongness judgments than for punishment judgment. This result provides some support for the claim that punishment judgments are more outcome-based than moral wrongness judgments (Cushman, 2008). As younger adults would by default consider to some extent outcomes when judging the punishability of an accidental harmdoer, age-related differences in relying on outcomes information may be reduced.

In the child development literature, recent attempts to account for the outcome-to-intent shift emphasized the role of ancillary changes occurring outside the moral domain such as in executive functioning and general cognitive abilities (Margoni & Surian, 2016), and in the child's tendency to attribute negligence to accidental harmdoers (Nobes et al., 2017). In our study we predicted age-related differences between younger and older adults, as we noticed that the general cognitive abilities and executive function components that are implied in processing

morally-relevant information (e.g., inhibitory control, working memory), and that need to be fully developed before a child can show intent-based judgments, are also the ones that decline in old age.

However, in the current study we did not find the predicted relationship between age group, working memory, and moral judgment—the effect of age group on moral judgment was not associated with working memory skills. As a tentative account of this null result, we may note that while working memory performance declines in old age (Hultsch, Hertzog, Small, McDonald-Miszczak, & Dixon, 1992; Park et al., 2002; Salthouse & Meinz, 1995), individual differences in this component of executive function may not be helpful in explaining age-related differences in moral judgment. It is possible that when working memory skills decline beyond a certain threshold level, old people start to heavily rely on their past experience: Older participants, to compensate for their executive functioning decline, but irrespective of the precise level of decline, may have attributed negligence to the agents because in real-life people who cause harm are often negligent. Future studies can investigate this and the role of executive functioning skills in age-related differences by measuring additional components of executive function such as inhibitory control and set shifting. Another possibility that could be investigated in future studies is that the threshold for attributing negligence lowers with aging, independently of the decline in the executive functioning. Again, employing different measures of executive function would be useful in determining whether an effect in negligence attribution remains significant even after controlling for executive variables.

A further limitation of this study concerns its within-subject design. In particular, one may notice that each participant completed both Task 1 and Task 2, and Task 2 always followed Task 1. It is then possible that having an explicit reference to negligence in Task 1 influenced the

judgment of negligence in Task 2. This can be true and needs to be assessed in future studies. Nevertheless, the main focus of the current study was on age-related differences in the tendency to rely on explicit information about agents' negligence and these were clearly found in the responses given to Task 1.

Another factor that may have exerted an influence on participants' ratings of moral wrongness may be the question wording: We asked participants to judge the wrongness of the action rather than how wrong (bad) was the character who performed the action. However, Nobes et al. (2016) found that adults' judgments of actions are more outcome-based than judgments of characters. Future studies may thus examine whether asking about characters rather than actions would help older adults to generate intent-based moral judgments, and thus reduce the age effect reported in the present study.

A final limitation concerns the study cross-sectional design. It can be suggested that the age-related differences we found may reflect a cohort effect. For example, it may be argued that older adults belong to a generation that, compared to the one of younger adults, gave less exculpatory value to the absence of negligence, perhaps due to a stricter education that set higher standards of carefulness. Future longitudinal studies could help decide whether the age-related differences in the tendency to rely on negligence information and, more in general, the intent-to-outcome shift, reflect a developmental or a cohort effect.

In conclusion, the present findings show that attributions of negligence play a crucial role in explaining age-related differences in moral judgment. Specifically, older adults were more prone to infer negligence from negative outcomes (Task 2), and this higher proneness can explain why older adults also condemned agents who were explicitly described as non-negligent (Task 1); it could be that older adults based their moral evaluations on their own inference that

the agents were negligent. This, in turn, can reflect a difficulty in suppressing assumptions of negligence or higher standards of carefulness. Moreover, it remains an open question whether (a) it is the focus on outcomes that prompts older adults to attribute negligence and subsequently condemn the accidental harmdoer; or (b) the higher tendency to attribute negligence can explain the intent-to-outcome shift. Regardless of the mechanism driving these age-related differences (deterioration in executive functions, higher reliance on experience), the present results point out the need to adopt a life-span view in studying moral development and to systematically investigate the often neglected developmental changes occurring in old age.

Acknowledgements

We thank XXX for helpful discussion and the students (Fiorenza Zancanaro and Simona Sardella) who helped collecting the data.

References

Alicke, M. D. (1992). Culpable causation. *Journal of Personality and Social Psychology, 63,* 368-378. https://doi.org/10.1037//0022-3514.63.3.368

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

Anderson, S. F., Kelley, K., & Maxwell, S. E. (2017). Sample-size planning for more accurate statistical power: A method adjusting sample effect sizes for publication bias and uncertainty. *Psychological science*, *28*, 1547-1562. https://doi.org/10.1177/0956797617723724

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

Buon, M., Jacob, P., Loissel, E., & Dupoux, E. (2013). A non-mentalistic cause-based heuristic in human social evaluations. *Cognition*, *126*, 149–155. https://doi.org/10.1016/j.cognition.2012.09.006

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. http://dx.doi.org/10.3758/s13423-016-1042-5

Chen, Y., & Blanchard-Fields, F. (2000). Unwanted thought: Age differences in the correction of social judgments. *Psychology of Aging, 15,* 475-482. https://doi.org/10.1037//0882-7974.15.3.475

Cushman, F. A. (2008). Crime and punishment: Distinguishing the role of causal and intentional analyses in moral judgment. *Cognition, 108,* 353-380. https://doi.org/10.1016/j.cognition.2008.03.006

Cushman, F. (2013). Action, outcome, and value: A dual-system framework for morality. *Personality and Social Psychology Review, 17(3),* 273-292. https://doi.org/10.1177/1088868313495594

Cushman, F., Sheketoff, R., Wharton, S., & Carey, S. (2013). The development of intent-based moral judgment. *Cognition, 127,* 6–21. https://doi.org/10.1016/j.cognition.2012.11.008

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior, 19,* 450-466. https://doi.org/10.1016/s0022-5371(80)90312-6

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. https://doi.org/10.1016/0022-1031(92)90037-k

Folstein, M., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189-198. https://doi.org/10.1037/t07757-000

Garon, N., Bryson, S., & Smith, I. (2008). Executive function in preschoolers: A review using an integrative framework. *Psychological Bulletin, 134,* 31-60. https://doi.org/10.1037/0033-2909.134.1.31

Hamlin, J. K. (2013). Failed attempts to help and harm: Intention versus outcome in preverbal infants' social evaluations. *Cognition, 128,* 451–474. https://doi.org/10.1016/j.cognition.2013.04.004

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging*, *28*, 826–839. https://doi.org/10.1037/a0030677

Hultsch, D. F., Hertzog, C., Small, B. J., McDonald-Miszczak, L., & Dixon, R. A. (1992). Short-term longitudinal change in cognitive performance in later life. *Psychology and Aging*, *7*, 571-584. https://doi.org/10.1037//0882-7974.7.4.571

Ligneau-Herve, C., & Mullet, E. (2005). Perspective-taking judgments among young adults, middle-aged, and elderly people. *Journal of Experimental Psychology: Applied,* 11, 53-60. https://doi.org/10.1037/1076-898x.11.1.53

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. https://doi.org/10.1027/1618-3169/a000395

Margoni, F., & Surian, L. (2016). Explaining the U-shaped development of intent-based moral judgments. *Frontiers in Psychology, 7,* 219. https://doi.org/10.3389/fpsyg.2016.00219

Margoni, F., & Surian, L. (2017). Children's intention-based moral judgments of helping agents. *Cognitive Development, 41,* 46-64. https://doi.org/10.1016/j.cogdev.2016.12.001

Margoni, F., & Surian, L. (2018). Infants' evaluation of prosocial and antisocial agents: A meta-analysis. *Developmental Psychology, 54,* 1445-1455. http://dx.doi.org/10.1037/dev0000538

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. https://doi.org/10.1037/xge0000234

Moran, J. M., Jolly, E., & Mitchell, J. P. (2012). Social-cognitive deficits in normal aging. *The Journal of Neuroscience, 32,* 5553–5561. https://doi.org/10.1523/jneurosci.5511-11.2012

Nobes, G., Panagiotaki, G., & Bartholomew, K. J. (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. 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. 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://dx.doi.org/10.1016/j.jecp.2009.08.001

Pazzaglia, F., Palladino, P., & De Beni, R. (2000). Presentazione di uno strumento per la valutazione della memoria di lavoro verbale e sua relazione con i disturbi della comprensione. *Psicologia Clinica dello Sviluppo*, *4*, 465-486.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and aging*, *17*, 299-320. https://doi.org/10.1037//0882-7974.17.2.299

Piaget, J. (1932). *The moral judgment of the child.* London: Kegan Paul.

Pratt, M. W., Diessner, R., Pratt, A., Hunsberger, B, & Pancer, S. M. (1996). Moral and social reasoning and perspective taking in later life: A longitudinal study. *Psychology and Aging*, *11*, 66–73. https://doi.org/10.1037//0882-7974.11.1.66

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior research methods*, *40*, 879-891. https://doi.org/10.3758/brm.40.3.879

Reuter-Lorenz, P. A., & Sylvester, C. Y. (2005). The cognitive neuroscience of aging and working memory. In: Cabeza R, Nyberg L, Park D, editors. *The cognitive neuroscience of aging* (pp. 186-217). New York: Oxford University Press.

Salthouse, T. A. (2004). What and when of cognitive aging. *Current Directions in Psychological Science*, *13*, 140–144. https://doi.org/10.1111/j.0963-7214.2004.00293.x

Salthouse, T. A., & Meinz, E. J. (1995). Aging, inhibition, working memory, and speed. *Journal of Gerontology: Psychological Sciences, 50,* 297-306. https://doi.org/10.1093/geronb/50b.6.p297

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. https://doi.org/10.1111/j.2044-835x.1983.tb00546.x

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/10.1037/h0080138

Shultz, T. R., Wright, K., & Schleifer, M. (1986). Assignment of moral responsibility and punishment. *Child Development*, *57*, 177-184. https://doi.org/10.2307/1130649

Siegal, M., & Peterson, C. C. (1998). Preschoolers' understanding of lies and innocent and negligent mistakes. *Developmental Psychology*, *34*, 332-342. https://doi.org/10.1037/0012-1649.34.2.332

Slaughter, V. (2015). Theory of Mind in infants and young children: A review. *Australian Psychologist*, *50,* 169-172. https://doi.org/10.1111/ap.12080

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. https://doi.org/10.1016/j.cognition.2017.06.029

Supplementary Material – Complete battery of scenarios

(The questionnaires were presented in Italian. Below we provide English translations of the instructions and scenarios.)

Instructions

Dear participant,

We thank you for your time. The purpose of the present study is to understand how people evaluate moral cases. The more general aim is to understand the psychological mechanisms that underlie human thinking. Below, we will ask you to read some instructions and then to read a number of scenarios and answer to the associated questions. The questionnaire will take approximately fifteen minutes to complete. If, after reading the instructions, you have doubts as to what you are supposed to do, please do not hesitate to ask for additional information.

Please read carefully the following instructions:

Your participation is on a voluntary basis and all the information we collect will be treated confidentially. You are free to interrupt the task at any time and for any reason. If you choose to participate, we ask you to read the stories and questions carefully, otherwise the data may not be reliable for the purposes of our research. In addition, we ask you to read the stories in the order in which they are presented, and answer to the questions in the order in which they appear. After reading a story carefully and answering to its related questions, please turn the page and continue with the next story. Please proceed in this way. Please provide your answer to a particular question next to it, by choosing a value between 0 and 10.

Task 1

First harm scenario – spinach

All-bad case (bad-intention/bad-outcome): Simon is grocery shopping for his grandmother who adores spinach. Recently there had been bacterial contamination of bagged spinach. At the market, Simon sees some bagged spinach on sale.

He thinks that bagged spinach may still be contaminated because of an incident just that day in his town.

Bagged spinach has been restocked at many markets, but some inspections aren't thorough and contaminated batches are missed. Simon buys his grandmother the spinach, and she cooks it, ending up in the hospital, violently ill.

All-neutral case (neutral-intention/neutral-outcome): Simon is grocery shopping for his grandmother who adores spinach. Recently there had been bacterial contamination of bagged spinach. At the market, Simon sees some bagged spinach on sale.

He thinks that bagged spinach is perfectly safe now because someone told him so.

It is safe to eat bagged spinach because it is no longer contaminated, in fact bagged spinach has been restocked at many markets. Simon buys his grandmother the spinach, and she cooks it. The meal is healthy and delicious.

Negligence accidental case (neutral-intention/negligence/bad-outcome): Simon is grocery shopping for his grandmother who adores spinach. Recently there had been bacterial contamination of bagged spinach. At the market, Simon sees some bagged spinach on sale.

He thinks that bagged spinach is perfectly safe now. Even though he came across the news of the contamination, he did not actively seek information about whether inspections on batches have been conducted.

Some inspections, indeed, were not thorough and contaminated batches are missed. Simon, without checking for it, buys his grandmother the spinach. She cooks it, ending up in the hospital, violently ill.

No-negligence accidental case (neutral-intention/no-negligence/bad-outcome): Simon is grocery shopping for his grandmother who adores spinach. Recently there had been bacterial contamination of bagged spinach. At the market, Simon sees some bagged spinach on sale.

He thinks that bagged spinach is perfectly safe now. Indeed, before going to the market, he checked on the newspaper whether all the contaminated batches have been withdrawn from the market.

However, some inspections were not thorough and contaminated batches are missed. Simon, thinking that it is not dangerous, buys his grandmother the spinach. She cooks it, ending up in the hospital, violently ill.

Second harm scenario – the dog

All-bad case (bad-intention/bad-outcome): Chloe works at the pound. Several new dogs have just come in. A lady comes in, interested in taking one dog home with her.

Chloe doesn't see the dogs, but her colleagues inform her that the new dogs all failed the health inspection and are infected with rabies.

The dogs are sick with rabies and will make their owners sick too by biting them. Chloe gives the lady one of the new dogs. It is infected with rabies and bites the lady on the very first day.

All-neutral case (neutral-intention/neutral-outcome): Chloe works at the pound. Several new dogs have just come in. A lady comes in, interested in taking one dog home with her.

Chloe doesn't see the dogs, but her colleagues inform her that the new dogs have been through a health inspection and will make good pets.

The dogs are healthy and active. Chloe gives the lady one of the new dogs. It is health, and the lady bonds immediately with it.

Negligence accidental case (neutral-intention/negligence/bad-outcome): Chloe works at the pound. Several new dogs have just come in. A lady comes in, interested in taking one dog home with her.

Chloe doesn't see the dogs, and because she is late for the birthday party of her son, she did not talk with her colleagues: She took for granted that the new dogs have been through a health inspection and will make good pets.

However, contrarily to what Chloe thinks, the dogs are sick with rabies and will make their owners sick too by biting them. Chloe gives the lady one of the new dogs. The dog is infected with rabies and bites the lady on the very first day.

No-negligence accidental case (neutral-intention/no-negligence/bad-outcome): Chloe works at the pound. Several new dogs have just come in. A lady comes in, interested in taking one dog home with her.

Chloe does see the dogs, and her colleagues reassure her that the new dogs have been through a health inspection and will make good pets.

However, contrarily to what Chloe thinks, the dogs are sick with rabies and will make their owners sick too by biting them. Chloe gives the lady one of the new dogs. The dog is infected with rabies and bites the lady on the very first day.

Third harm scenario – the zoo

All-bad case (bad-intention/bad-outcome): Robert is at the zoo with his nephew. They are watching the dolphin show when the nephew complains that his stomach hurts.

Robert thinks that his nephew's stomach hurts because of a major surgical operation he had several weeks ago; Robert thinks that he needs medical attention immediately.

The nephew is really sick. After the recent operation, stomach pain could indicate really serious complications. Robert takes him to see the monkeys although he thinks that his nephew is really sick. His nephew starts feeling worse and soon blacks out because of severe internal bleeding.

All-neutral case (neutral-intention/neutral-outcome): Robert is at the zoo with his nephew. They are watching the dolphin show when the nephew complains that his stomach hurts.

Robert thinks that his nephew's stomach hurts because he ate too much cotton candy and fried dough that afternoon, and he doesn't know that the nephew has recently undergone a major surgical operation; Robert thinks his nephew just needs to walk it off.

The nephew is really fine. His stomach sometimes hurts when he eats too much, but he usually feels better after an hour or so. Robert takes him to see the monkeys. His nephew starts feeling better, and they see nearly all the zoo exhibits.

Negligence accidental case (neutral-intention/negligence/bad-outcome): Robert is at the zoo with his nephew. They are watching the dolphin show when the nephew complains that his stomach hurts.

Robert thinks that his nephew's stomach hurts because he ate too much cotton candy and fried dough that afternoon, and he doesn't know that the nephew has recently undergone a major surgical operation. The uncle does not care much about his nephew, and, after the first nephew's manifestations of pain, he does not call the parents to make sure the child does not suffer from any medical condition. Robert thinks his nephew just needs to walk it off.

However, the nephew is really sick. After the recent operation, stomach pain could indicate really serious complications. Robert takes him to see the monkeys. His nephew starts feeling worse and soon blacks out because of severe internal bleeding.

No-negligence accidental case (neutral-intention/no-negligence/bad-outcome): Robert is at the zoo with his nephew. They are watching the dolphin show when the nephew complains that his stomach hurts.

Robert thinks that his nephew's stomach hurts because he ate too much cotton candy and fried dough that afternoon, and he doesn't know that the nephew has recently undergone a major surgical operation. The uncle asks often about him to his parents, but the parents did not tell the uncle about the medical operation. Robert thinks his nephew just needs to walk it off.

However, the nephew is really sick. After the recent operation, stomach pain could indicate really serious complications. Robert, who did not know about the operation, takes the nephew to see the monkeys. The nephew starts feeling worse and soon blacks out because of severe internal bleeding.

Fourth harm scenario – jellyfish

All-bad case (bad-intention/bad-outcome): Joanna and one of her acquaintances are on a boat in a part of the sea with lots of jellyfish. Joanna's acquaintance asks her if she can go for a swim.

Since Joanna read that the local jellyfish are poisonous, she thinks it is not safe to swim in the sea.

It is not safe to swim in the sea because the jellyfish sting and their stings are poisonous. Joanna tells her acquaintance to go for a swim. Her acquaintance does, gets stung by jellyfish and goes into shock.

All-neutral case (neutral-intention/neutral-outcome): Joanna and one of her acquaintances are on a boat in a part of the sea with lots of jellyfish. Joanna's acquaintance asks her if she can go for a swim.

Since Joanna read that the local jellyfish are harmless, she thinks it is safe to swim in the sea.

It is perfectly safe to swim in the sea because the jellyfish don't sting and are harmless. Joanna tells her acquaintance to go for a swim. Her acquaintance does and enjoys the swim.

Negligence accidental case (neutral-intention/negligence/bad-outcome): Joanna and one of her acquaintances are on a boat in a part of the sea with lots of jellyfish. Joanna's acquaintance asks her if she can go for a swim.

Joanna thinks it is safe to swim in the sea, but she did not verify with anybody her belief. Before leaving with the boat, notwithstanding she was in charge of organizing the trip, she did not collect information about the sea conditions.

Contrarily to what Joanna thinks, it is not safe to swim in the sea because the jellyfish sting and their stings are poisonous. Joanna, answering without due care and without being informed, tells her acquaintance to go for a swim. Her acquaintance does, gets stung by jellyfish and goes into shock.

No-negligence accidental case (neutral-intention/no-negligence/bad-outcome): Joanna and one of her acquaintances are on a boat in a part of the sea with lots of jellyfish. Joanna's acquaintance asks her if she can go for a swim.

Joanna organized the trip with due care and read on an important newspaper that the local jellyfish are harmless. Therefore, she thinks it is safe to swim in the sea.

However, contrarily to what Joanna thinks, it is not safe to swim in the sea because the jellyfish sting and their stings are poisonous. Joanna tells her acquaintance to go for a swim since she thinks it is safe. Her acquaintance does, gets stung by jellyfish and goes into shock.

Task 2

First accidental harm case – spinach

Neutral-intention/bad-outcome: Simon is grocery shopping for his grandmother who adores spinach. Recently there had been bacterial contamination of bagged spinach. At the market, Simon sees some bagged spinach on sale.

He thinks that bagged spinach is perfectly safe now because someone told him so.

Bagged spinach has been restocked at many markets, but some inspections aren't thorough and contaminated batches are missed. Simon, thinking that it is not dangerous, buys his grandmother the spinach, and she cooks it, ending up in the hospital, violently ill.

Second accidental harm case – the dog

Neutral-intention/bad-outcome: Chloe works at the pound. Several new dogs have just come in. A lady comes in, interested in taking one dog home with her.

Chloe doesn't see the dogs, but her colleagues inform her that the new dogs have been through a health inspection and will make good pets.

However, contrarily to what Chloe thinks, the dogs are sick with rabies and will make their owners sick too by biting them. Chloe gives the lady one of the new dogs thinking that the dog is healthy. It is infected with rabies and bites the lady on the very first day.

Third accidental harm case – the zoo

Neutral-intention/bad-outcome: Robert is at the zoo with his nephew. They are watching the dolphin show when the nephew complains that his stomach hurts.

Robert thinks that his nephew's stomach hurts because he ate too much cotton candy and fried dough that afternoon, and he doesn't know that the nephew has recently undergone a major surgical operation; Robert thinks his nephew just needs to walk it off.

However, the nephew is really sick. After the recent operation, stomach pain could indicate really serious complications. Robert takes him to see the monkeys because he doesn't know about the operation. His nephew starts feeling worse and soon blacks out because of severe internal bleeding.

Fourth accidental harm case – jellyfish

Neutral-intention/bad-outcome: Joanna and one of her acquaintances are on a boat in a part of the sea with lots of jellyfish. Joanna's acquaintance asks her if she can go for a swim.

Since Joanna read that the local jellyfish are harmless, she thinks it is safe to swim in the sea.

However, contrarily to what Joanna thinks, it is not safe to swim in the sea because the jellyfish sting and their stings are poisonous. Joanna tells her acquaintance to go for a swim since she thinks it is safe. Her acquaintance does, gets stung by jellyfish and goes into shock.

Supplementary Material – Additional analyses

The effect of question order on participants' punishment judgments (Task 1)

Previous studies have shown that asking children and adults to judge an action (e.g., 'selling a sick dog') leads to outcome-based judgments, whereas asking participants to judge the character of an agent (e.g., Was the agent good or bad?) leads to intention-based judgments (see Nobes et al., 2016). Therefore, the wording of the moral wrongness question which we used in the current study might have led participants to focus more on the agent's action than on the agent's character, and this, in turn, may have prompted participants to judge based on outcome information. However, this issue does not apply to the wording of the punishment question, which was agent-focused. Furthermore, the presentation order of the questions may have impacted judgments due to carry-over effects. Specifically, having the punishment judgment first may lead participants to focus on the moral character of the person when responding to the subsequent moral wrongness question, while having the moral wrongness judgment first may prompt participants to focus on the action when responding to the punishment question. To examine this latter possibility, we reanalysed the punishment data testing for order effects.

We submitted the punishment judgments to a 2 (Age group: old vs. young) × 2 (Negligence level: present vs. absent) × 2 (Order of test question: wrongness judgments first vs. punishment judgments first) mixed-factor ANOVA, with repeated measures on negligence level (see Figure S1). The analysis revealed a significant main effect of negligence level, $F(1, 77) = 15.76$, $p < .001$, $\eta_p^2 = .17$, $f = 0.46$. Accidental harms that resulted from negligence received higher punishment ratings ($M = 6.74$, 95% CI [5.98, 7.50]) than accidental harms that did not result from negligence ($M = 4.63$, CI [3.79, 5.47]). Critically, this effect was qualified by a marginally significant Age group × Negligence level interaction, $F(1, 77) = 3.81$, $p = .055$, $\eta_p^2 =$

.05, $f = 0.23$. Pairwise comparisons showed that older participants were less influenced by negligence level ($M_{Present} = 7.38$, CI [6.30, 8.45]; $M_{Absent} = 6.30$, CI [5.11, 7.49]), $F(1, 77) = 2.03$, $p = .158$, $\eta_p^2 = .03$, $f = 0.16$, than younger participants ($M_{Present} = 6.11$, CI [5.04, 7.18]; $M_{Absent} = 2.96$, CI [1.78, 4.14]), $F(1, 77) = 17.62$, $p < .001$, $\eta_p^2 = .19$, $f = 0.48$. Furthermore, there was a significant main effect of age group, $F(1, 77) = 14.68$, $p < .001$, $\eta_p^2 = .16$, $f = 0.44$. Overall, older participants gave higher punishment ratings ($M = 6.84$, CI [5.99, 7.69]) than younger participants ($M = 4.53$, CI [3.69, 5.38]).

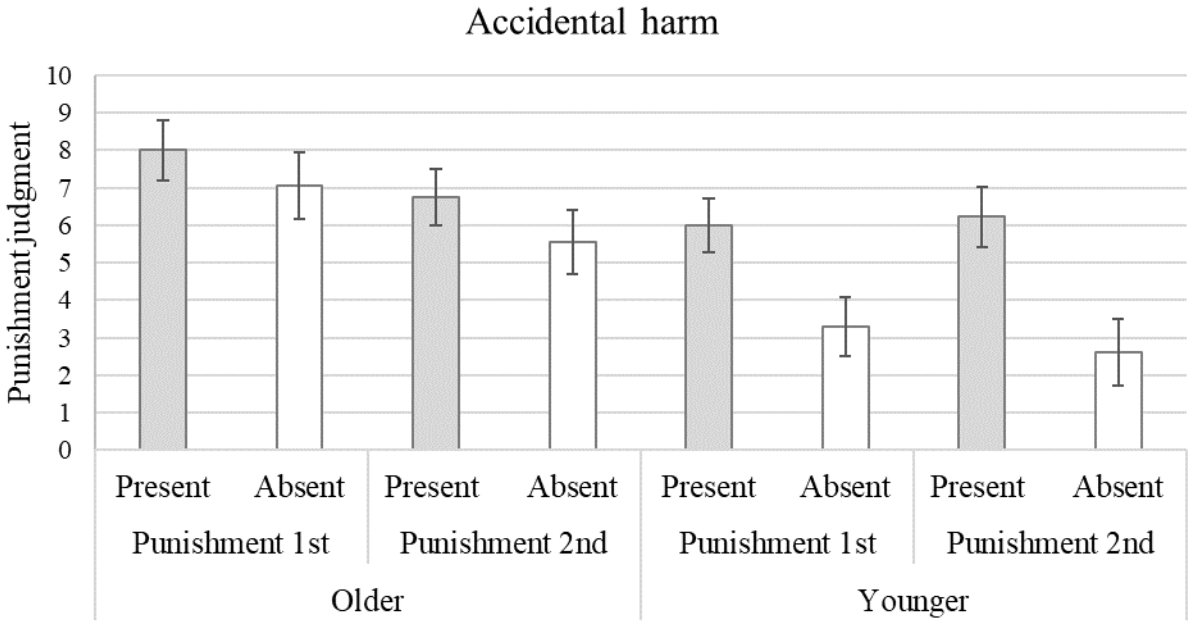

Figure S1. Average punishment judgments by age group condition (old vs. young), negligence level (present vs. absent) and order of judgments presentation (punishment ratings first vs. second). Error bars indicate standard errors of the mean.

Returning to the tests of interest, there was no significant main effect of order, $F(1, 77) = 1.80$, $p = .184$, $\eta_p^2 = .02$, $f = 0.14$. Importantly, there was no significant Age group × Order interaction, $F(1, 77) = 0.90$, $p = .346$, $\eta_p^2 = .01$, $f = 0.10$, no significant Order × Negligence level

interaction, $F(1, 77) < 0.30$, $p = .586$, $\eta_p^2 < .01$, $f < 0.10$, and no Age group × Order × Negligence level interaction, $F(1, 77) = 0.10$, $p = .756$, $\eta_p^2 < .01$, $f < 0.10$.

The effect of question order on participants' punishment judgments (Task 2)

We submitted the punishment judgments to a 2 (Age group: old vs. young) × 3 (Order of test question: moral wrongness/punishment/negligence [1] vs. punishment/negligence/moral wrongness [2] vs. negligence/moral wrongness/punishment [3]) ANOVA. The analysis revealed a significant main effect of age group, $F(1, 74) = 27.22$, $p < .001$, $\eta_p^2 = .27$, $f = 0.61$, and a significant main effect of order, $F(2, 74) = 3.89$, $p = .025$, $\eta_p^2 = .10$, $f = 0.32$. There was a marginally significant Age group × Order, $F(2, 74) = 2.53$, $p = .087$, $\eta_p^2 = .06$, $f = 0.25$. Pairwise comparisons showed that the effect of age group was only significant for the orders of test question presentation 2 and 3 (Fs > 10.70, ps < .003). Again, this effect was driven by older participant's judgments. In the group of older participants, there was a significant difference between order 1 ($M_{Order1} = 5.21$, CI [3.49, 6.94]) and order 2 ($M_{Order2} = 7.77$, CI [5.98, 9.56]), as well as between order 1 and order 3 ($M_{Order3} = 9.62$, CI [7.83, 11.40]). No other comparisons were statistically significant.