Current Psychology (2023) 42:6752–6767
https://doi.org/10.1007/s12144-021-01956-5

Plenty of blame to go around: Attributions of responsibility in a fatal autonomous vehicle accident

Christopher J. Copp[1] · Jean J. Cabell[1] · Markus Kemmelmeier[1]

Accepted: 1 June 2021 / Published online: 26 June 2021

Abstract

Autonomous vehicles (AV) promise a reduction in the number of deadly traffic accidents. However, should accidents occur, attributions of responsibility are complicated by the fact that there is a human agent (driver) and a non-human agent (AV), and thus responsibility is likely shared between parties. In two studies, participants ($n = 310$ and $n = 260$) read a vignette modeled after an actual lethal AV accident. Across four experimental conditions, participants were told that the human driver either needed to maintain oversight of the AV; did not need to maintain oversight of the AV; did not specify whether the human needed to maintain oversight of the AV; or the artificial intelligence was turned off and the human driver was fully in control. Participants assigned responsibility to the human driver, the AV company, the pedestrian, and an act of God, and determined whether the human driver and company CEO should be held criminally responsible in court. Consistent with previous research, the human driver was held most responsible regardless of oversight condition. However, companies were not absolved of responsibility, even when they required the human driver to maintain oversight of the AV. Implications of these findings for the introduction and legal regulation of AVs are discussed.

Keywords Autonomous vehicles · Self-driving cars · Attribution of responsibility · Artificial intelligence · Agency · Blame

The advent of artificial intelligence (AI) requires individuals to consider non-human agents when assigning blame and responsibility for accidents. One such case is that of autonomous vehicles (AV), which are increasingly seen on the streets of industrialized societies. Though fully autonomous vehicles are still rare, the number of vehicles with some level of autonomy is increasing. In many respects, this technological change is welcome. Humans, and especially human drivers, are without a doubt the most frequent cause of traffic accidents (National Highway Traffic Safety Administration, 2015), which combined cause nearly 40,000 fatalities every year on U.S. streets (National Safety Council, 2020). AV technology promises to reduce these deaths because AVs are much less prone to accidents due to technology created to prevent collisions.

Nevertheless, things can go wrong. There have been several accidents involving AVs, some of which were fatal. When determining a responsible party and placing blame, people are highly focused on agency; that is, an actor choosing a course of action. For most of history, humans were perceived to be the only source of agency.[1] This is especially true in the determination of blame. There is plenty of evidence that people selectively identify human agency as the cause of accidents, especially when people perceive an undesirable outcome that could have been prevented (Alicke, 2000). Consistent with this idea, people tend to be satisfied with human agency as a sufficient cause of an outcome when tracing the causes of an undesirable event to its origins, even when the thus identified human action was caused by other circumstances (e.g., Hilton et al., 2016).

In line with this emphasis on human agency in causal attribution, Awad et al. (2020) recently examined attributions of blame for AV accidents. The authors focused on different types of agency that exist when human drivers collaborate with AI in the operation of a vehicle. Specifically, the authors examined when a pedestrian was killed as a function of (a) the failure of human drivers to intervene in a malfunction of the AV; (b) the erroneous intervention of a human driver in the otherwise correct operation of an AV; (c) the failure of an AV

✉ Markus Kemmelmeier
markusk@unr.edu

[1] Interdisciplinary Social Psychology Ph.D. Program, University of Nevada, Reno, Mail Stop 1300, Reno, NV 89557-1300, USA

[1] Beyond humans, agency also tends to be attributed to deities or beings such as spirits, totems etc. Typically, such beings are represented in ways that attribute human characteristics to them.

to intervene and correct the errors of a human driver, and (d) the erroneous intervention of an AV in the otherwise correct driving of a human driver. Across five studies, observers were most inclined to lay blame and responsibility with the human driver in scenarios (a) and (c) that involved both "drivers" making errors, even when an AV performed an action identical to that of the human driver. This finding confirmed the well-documented readiness to focus on human agency as a causal force, even when AI may have presented itself as another source of agency. Broadly, Awad et al.'s (2020) findings support the idea of a "moral crumple zone" proposed by Elish (2019). According to Elish, when faced with the challenge of having to attribute responsibility to either an automated system or human actors, people are disproportionately inclined to hold the system faultless, and blame the human operators even when they have limited control over the situation. In this sense, humans serve as moral buffers against blaming technology and its manufacturers.

Awad et al.'s (2020) results are also broadly consistent with research by McManus and Rutchick (2019). In their study, McManus and Rutchick had participants read about a scenario in which an AV accident resulted in the death of a driver, or conversely, of two construction workers. The authors varied whether the human was driving a traditional vehicle or driving an AV that had either been programmed by the vehicle manufacturer or by the owner himself. Not surprisingly, when the AV was in full control, participants attributed less blame to the human owner, unless he programmed the AV himself. This finding highlights that perceived agency (whether the human or the robot was in charge) was the prime determinant of blame (cf. Bennett et al., 2020).[2] However, the scenarios used by McManus and Rutchick focused on the special case in which death was unavoidable such that either the driver or others were killed (see also Bonnefon et al., 2016). Such scenarios are useful to evaluate existential choice as epitomized by the frequently employed trolley problem (Foot, 1967; but see De Freitas et al., 2020). Yet, most fatal accidents involving AVs do not involve an unavoidable death. In many vehicle accidents the harm is asymmetric, such as when a vehicle collides with a pedestrian or a cyclist, in which greater harm to the latter party is expected regardless of who caused the accident.

The work by McManus and Rutchick (2019) also leaves several issues unresolved. Most prominently, at the present time all AVs on streets in the U.S., Australia, and many other countries require that humans maintain continuous oversight and are ready to intervene in the operations of the AV at any time (e.g., Arizona's Executive Order 9, 2015, or Nevada Revised Statue 482A; National Transport Commission, 2020). The work by McManus and Rutchick (2019) did not make any mention of this essential legal obligation of continuous human oversight when operating AVs.

Awad et al.'s (2020) research is more realistic in that it examined the asymmetrical nature of most accidents, namely by focusing on a situation in which an AV kills a pedestrian in an accident that could have been avoided. However, because their experimental scenario is constructed in a way that highlights that there are often two drivers (e.g., a human and a machine), it remains unclear how requirements of AV oversight may modify attributions of blame. Moreover, the authors did not report attributions to other parties outside of the human driver and the machine, and it is not clear to what extent research participants might have blamed other parties or circumstances (e.g., Pöllänen et al., 2020).

In our work, we focus on the first and thus far only fatal AV accident involving a pedestrian which occurred on March 18th, 2018 in Tempe, AZ. A little before 10:00 p.m., Elaine Herzberg, 49, was walking her bicycle across a dark street outside of a pedestrian crosswalk. An autonomous Volvo owned and operated by Uber collided with Herzberg, resulting in her death. Although the tragedy of this event was transparent, the responsibility for the accident was much more opaque. The vehicle's driver, Rafaela Vasquez, 44, who was hired by Uber to oversee the vehicle in the event of a malfunction, was distracted. According to Tempe police and an investigation by the National Traffic Safety Board (NTSB), Rafaela had often taken her eyes off of the road for extended periods of time (Randazzo, 2019). Further, she was actively streaming the television show "The Voice" to her phone, and was observed looking at her phone immediately preceding the accident by video taken from inside the vehicle (BBC News, 2019).

There was also evidence to implicate the pedestrian herself in the accident. A toxicology report determined that Herzberg had methamphetamine and marijuana in her system, and may have been intoxicated at the time of the incident (Randazzo, 2019). Further, had Herzberg been crossing at a designated crosswalk area, the AV might have been quicker to identify her as a pedestrian and the collision might have been avoided (McCausland, 2019).[3] However, the victim's family is suing the city of Tempe for $10 million, stating that the paving in the median was indicative of a legal road crossing.

The programming of the Uber vehicle was also implicated. According to the same NTSB report, the vehicle lacked "the capability to classify an object as a pedestrian unless that object was near a crosswalk" (McCausland, 2019). Despite a series of sensor systems including lidar, radar, and traditional cameras that were engineered to identify objects in the road and to predict the path of those objects, the vehicle failed to

[2] McManus and Rutchik (2019) also tested to what extent the availability of an override switch shaped blame attributions. If the human owner could have overridden the actions of the AV, the driver was not blamed any more or less compared to when he was operating a manual vehicle. However, more blame was assigned when the driver did override the actions of the vehicle. The latter finding highlighted commission as a critical element in blame attributions.

[3] The only way for the vehicle to "intervene" would have been for it to notify the driver (Vasquez) because the automated brake function had been disabled.

correctly identify Herzberg as she crossed the street (Gonzales, 2019). Although the sensors detected Herzberg as an object six seconds before the impact, they did not correctly identify her as a pedestrian nor determine her trajectory. About 1.3 s before the impact, the vehicle determined that braking was necessary, but Uber had disabled the emergency braking system and instead was reliant upon the safety driver to stop the vehicle in the event of a malfunction (McCausland, 2019).

The ambiguity of this case underscores the potential problem in the ubiquitous adoption of AVs. When an accident occurs in a traditional vehicle, the accountable party is typically straightforward. For example, were a driver to have caused an accident that resulted in the death of a pedestrian while being in reckless disregard of the road, he or she would likely have faced criminal charges. However, as illustrated in the case of Elaine Herzberg, AV accidents likely have multiple points of failure rather than just one. Further, system failures in AVs are difficult to attribute to a lone actor, and are instead diffuse throughout the manufacturing organizations.

Within a year of the accident, Arizona prosecutors decided that Uber would not be held criminally liable for the death of Elaine Herzberg (Wamsley, 2019). More than two and a half years after the incident, charges of negligent homicide were brought against Rafaela Vasquez (Elish, 2020).[4] As of this writing, her trial is set for August 10, 2021.

To examine attributions of blame and responsibility with regard to accidents involving AVs, we created vignettes that were inspired by the case of Elaine Herzberg, though the scenario was slightly simplified. For instance, in our scenario we referred to the manufacturer and the company operating the vehicle as the same company, though in reality these could differ. We also took some liberties with regard to the driver, but emphasized her experience as a professional driver. Lastly, we highlighted that the pedestrian did not cross at a designated crosswalk (see Method for our complete vignette).

In our research, we experimentally varied whether the driver was required to oversee the vehicle at all times, as was the case in the accident involving Elaine Herzberg; whether the driver did not need to maintain oversight at all times; or whether the AI enabling autonomous driving had been turned off. In a fourth and final condition, we did not state whether or what kind of oversight was required. Consistent with established models of responsibility attribution (Shaver, 2012; Weiner, 1995), we anticipated that the more control the driver had, the more responsibility for the accident would be attributed to her. That is, less responsibility should be attributed to the driver when the AV did not require any oversight than when the AV required oversight or when the autonomous driving capability was turned off. Conversely, less responsibility should be attributed to the manufacturer when the driver had more control.

Note that these three levels of human control resembled those modeled by Pöllänen et al. (2020), who varied whether the vehicle was manually driven, semi-automatic, or fully automatic; however, the definition of "semi-automatic" given to their participants emphasized that the AV would request intervention from the human driver. In cases of emergency, as occurred in the accident involving Elaine Herzberg, this scenario does not sufficiently highlight what is currently a legal requirement: namely, that the human driver must oversee the AV and intervene when the situation requires it, regardless of whether he or she is prompted by the vehicle to do so (see, e.g., Arizona's Executive Order 9, 2015, or Nevada Revised Statue 482A n.d.). Hence, our scenario is arguably more realistic.

Despite the findings by Awad et al. (2020), we did not necessarily assume that the driver would be blamed the most, because in a fully autonomous vehicle, agency rests primarily with the vehicle and there is no necessity for the human driver to pay attention to the road. Hence, we expected that the company would be held more accountable than the driver when the AV is fully autonomous. Given the reality of the Herzberg accident, we also included the pedestrian as a potential target for attribution (see also Bennett et al., 2020; Pöllänen et al., 2020) to expand upon the scenarios from Awad et al. (2020) and McManus and Rutchick (2019). In addition, we also offered participants the option of designating no culpable party or describing the event as an act of God, that is, an event that is defined in law as being outside of human control.[5]

The present research also investigates to what extent attributions of responsibility and blame are qualified by the type of corporate defendant. Research has demonstrated that in civil court, corporate defendants are treated more punitively compared to individual defendants (MacCoun, 1996). This disparity is in some part due to the assumption that corporations tend to have "deeper pockets" and are more likely to pay higher compensatory and punitive damages (Hans, 2000; Vidmar, 1999). At the present time, companies engaged in developing autonomous mobility vary dramatically in size, ranging from multi-billion dollar companies, such as Google's Waymo (1500 employees), to relatively small companies focusing on specialized application, such Locomation.ai (20 employees). In this research we hypothesized that, as in prior research, larger, wealthier companies would be held more responsible.

[4] The present research was conducted in Spring of 2020, when Rafaela Vasquez had not been charged yet, and when it appeared that she would not be charged. The present research was inspired in part by the question of why she had not been charged, even when existing research suggested that the human driver would be most likely identified as the party most responsible.

[5] The present studies also explored various moderator variables, including self-reported knowledge of and attitudes toward AVs, as well as political orientation, as existing research suggests that conservatives tend to be more pro-business (e.g., Heinze et al., 2014), and are more likely than liberals to embrace an ideology of personal responsibility (Carroll et al., 1987; Schlenker et al., 2012). However, none of the variables yielded reliable and consistent effects across our two studies; hence, they are not reported further.

Study 1

Method

Participants

We recruited 350 participants from Amazon's Mechanical Turk (MTurk) and compensated them $0.75. Our final sample was 310 participants after removing 34 participants with duplicate IP addresses, four participants who took fewer than 60 s to complete the survey (completion time $M = 334$ s, $Md = 216$ s), and two participants who agreed to participate but did not answer any items. The sample self-identified as 60.0% men, and included 70.3% Caucasians/Whites, 10.3% African Americans/Blacks, 9.0% Latin/Hispanic, 4.5% Asians. The mean age was 36.0 ($SD = 11.3$; range = 19–72).

Design

We randomly assigned participants to a 4 (oversight: required vs. not required vs. not mentioned vs. AI not engaged) × 2 (company size: small vs. large) between-subjects factorial design. In all conditions, the participant read a scenario about Kate, a test driver for an automobile company entering the AV market. Kate had positive performance reviews and experience behind the wheel of AVs. While piloting one of the vehicles, she was looking at her phone when a pedestrian crossed the street illegally, killing the pedestrian instantly.

Materials

Vignette All participants read a version of the following vignette (with text varied between conditions underlined):

> "Kate is 30 years old, and has been working as a test driver for a small and upcoming automobile company for the last two years. This automobile company is attempting to enter the autonomous vehicle, or self-driving car, market in the near future. She has always received positive performance reviews from the company.
> The automobile manufacturer has stated that their self-driving car technology needs consistent human oversight, with a driver paying attention and being able to intervene at all times.
> Kate has thus far logged over 10,000 miles of experience behind the wheel of self-driving vehicles without incident.
> Last week, Kate was piloting the vehicle with the self-driving mode of the car turned on. She was looking at her phone and did not see a pedestrian who was crossing the street illegally, not at an approved crosswalk. The car struck the pedestrian, killing them instantly."

For half of our participants, the company for which the driver (Kate) worked was "small and upcoming," whereas for the other half it was "large and established." In the *oversight required* condition participants learned that "their self-driving car technology needs consistent human oversight, with a driver paying attention and being able to intervene at all times." In the *oversight not required* condition participants read that "their self-driving car technology was in the final phase of development and does not need consistent human oversight." In the *AI not engaged* condition, participants read a sentence simply stating that she "was piloting the vehicle with the self-driving mode of the car turned off." Lastly, in the *not mentioned* condition, participants did not receive any information about whether oversight was required or not, though as in the vignette above, the vehicle was described as "self-driving."[6]

Responsibility To examine responsibility for the accident, participants rated how responsible Kate, the automobile manufacturer, and the pedestrian were on a scale from 0 (*not at all responsible*) to 100 (*completely responsible*). They also rated the extent to which they believed no one was responsible from 0 (*disagree*) to 100 (*agree*); specifically, the item to which participants responded was "No one is responsible for the accident, bad things just happen sometimes." This amounts to what is legally referred to as an "act of God," a phrase we use throughout the remainder of the manuscript. The response scale was an online slider scale, with the slider initially being positioned at 50 (midpoint) and requiring participants to drag it to the desired position on the scale.

Criminal Responsibility Participants answered a yes or no question as to whether they believed the driver and the CEO of the company should be held criminally responsible for the accident.

Demographic Questions Participants answered demographic questions assessing their age, gender, race and ethnicity, level of education, and political affiliation (the latter measured on a scale from 1 [*more liberal*] to 5 [*more conservative*]).[7]

Procedure

After reading and agreeing to a consent sheet, participants were instructed to read a story about a female driver named Kate. Participants were then randomly assigned to read one of

[6] Some conditions in our study do correspond to specific automation levels as stipulated by the Society of Automotive Engineers (SAE). Our "AI not engaged" condition corresponds to level 0 (no automation). Our "oversight not required" condition suggests level 5 (human attention not required). Our "oversight required" leaves it open if this is level 3 or 4 (both require human interaction in case of error or system failure).

[7] In both studies reported here, participants also responded to questions pertaining to monetary liability, and knowledge and trust of autonomous vehicles, and trust in technology in general. These variables were either considered problematic or did not yield any meaningful results. Additional information about these findings can be obtained from the authors.

eight vignettes. Next, participants answered the responsibility, criminal responsibility, AV and technology knowledge and trust, and demographic questionnaire.

Results

Responsibility

To account for the nested nature of the responsibility measures, we used a linear mixed effects model to examine the effects of oversight and company size on participants' responsibility judgments. The fixed effects included oversight, company size, the party responsibility rating (i.e., driver, company, pedestrian, act of God), and all interactions between these factors, whereas the random effects included participants. Analyses were conducted using the *lme4* package in R (Bates et al., 2015).

The multilevel model results indicated a significant main effect of company size, $F(1, 302) = 8.81$, $p = .003$. Contrary to our prediction, participants attributed more responsibility when the car company was *small* ($M = 55.10$, $SD = 32.15$) compared to *large* ($M = 48.96$, $SD = 34.76$). However, company size was not involved in any interaction effect. Overall, this finding did not confirm our hypothesis concerning the "deep pockets" effect.

Importantly, there was also a main effect of the party responsibility rating $F(3, 906) = 76.90$, $p < .001$, such that participants attributed higher levels of responsibility to the driver compared to the company, pedestrian, and to the act of God option, all $ps < 0.001$; see Table 1 for a summary of means). Critically, this effect was qualified by a significant interaction between oversight and the party responsibility rating $F(9, 906) = 5.35$, $p < 0.001$ (see Fig. 1).

Most Held Responsible Not surprisingly, the driver was always considered the party most responsible when either the *AI was not engaged* or when *oversight was required*, with means substantially greater than those for all other parties, all $ps < .001$; see Table 1 and Fig. 1. However, when *oversight was not required*, there was no difference in responsibility attributions between the driver and the company, $p = .55$, with both being considered more responsible than the pedestrian or the act of God option, all $ps < .043$. Importantly, the driver was not relieved of responsibility when *oversight was not required*. When *oversight was not mentioned*, there was no difference in how much responsibility was attributed to the driver, the company, or the pedestrian, all $ps > .15$, though each was considered more responsible than an act of God, all $ps < .001$. (For ease of comprehension, the same Study 1 data as summarized in Table 1 and Fig. 1 are also displayed in Appendix Part A, allowing for a direct comparison between attributions of responsibility to different sources of agency.)

Driver Responsibility As could be expected, participants attributed more responsibility to the driver when the *AI was not engaged* ($M = 74.22$, $SD = 25.65$) than when *oversight was not required* and when *oversight was not mentioned*, $ps < .021$. Pairwise analyses indicated that participants attributed more responsibility to the driver when *oversight was required* ($M = 74.20$, $SD = 24.72$) than when *oversight was not required* ($M = 61.28$, $SD = 27.59$), $p = .010$, and when *oversight was not mentioned* ($M = 62.73$, $SD = 25.65$; $p = .020$). No differences emerged between *oversight not required* and *oversight not mentioned*, indicating that participants considered drivers to be equally responsible across these conditions. Similarly, there were also no significant differences between responsibility in the *AI not engaged* and *oversight required* conditions.

Company Responsibility As hypothesized, the company was considered more responsible when *oversight was not required* ($M = 64.17$, $SD = 25.70$) compared to when *oversight was required* ($M = 51.60$, $SD = 32.38$), $p = .010$. Participants also attributed more responsibility when *oversight was not mentioned* ($M = 59.93$, $SD = 30.12$), $p < .001$, *oversight was not required*, $p < .001$, or when *oversight was required*, $p = .012$, than when the *AI was not engaged* ($M = 39.51$, $SD = 37.66$). These findings suggest that companies are still held responsible even when they require driver oversight, though to a lesser extent than when no oversight is required.

Act of God and Pedestrian Participants were more likely to consider the accident an act of God when *oversight was not required* ($M = 38.26$, $SD = 32.75$), $p = .009$, or *oversight was not mentioned* ($M = 36.84$, $SD = 33.27$), $p = .027$, in comparison to when the *AI was not engaged* ($M = 25.62$, $SD = 31.77$). The mean for oversight required fell in-between the other three, without being significantly different from them, all $p > .098$. Attributions of responsibility to the pedestrian did not differ by levels of oversight.

Criminal Liability

Participants were asked whether they believed the driver or the CEO of the AV manufacturer should be held criminally liable for the death of the pedestrian. A logistic regression model was used to test whether the primary manipulation of oversight, company size, and their interaction was related to criminal responsibility.

Driver A total of 52% of our sample believed that Kate, the driver, should be held criminally responsible. There was no significant main effect of company size or the interaction effect between oversight condition and company size ($ps > .18$). A significant main effect for the oversight condition, $\chi^2(3) = 8.73$, $p = .03$, showed that participants were most likely to believe that the driver should be held criminally responsible when the *AI was not engaged* (63%) compared to when *oversight was not required* (45%, $p = .026$) and when *oversight was not mentioned* (42%, $p = .011$), though no reliable difference with the *oversight required* condition emerged (56%, $p = .38$). Interestingly, when the AI was

Table 1 Perceived responsibility and criminal liability (Study 1 & Study 2)

	Driver		Company		Pedestrian		Act of God	
	M	(*SD*)	*M*	(*SD*)	*M*	(*SD*)	*M*	(*SD*)
Responsibility rating								
Study 1	68.25[a]	(26.49)	53.64[b]	(33.04)	52.73[b]	(32.48)	33.50[c]	(32.70)
Study 2	59.78[a]	(26.30)	52.03[b]	(30.43)	51.38[b]	(29.40)	33.07[c]	(34.74)
Criminal liability								
Study 1	52.13%		38.36%					
Study 2	58.46%		47.69%					

Responsibility ratings were recorded on a 0–100 scale; criminal liability was recorded on a percent scale

engaged, there was no difference between whether *oversight was required* or *not required*, $p = .17$. No other effects were significant. Driver criminal responsibility was positively correlated with general driver responsibility ($r = 0.48$, $p < .001$), suggesting participants' responsibility attributions are related to their beliefs regarding criminal responsibility.

Company CEO Approximately 38% of our sample believed that the CEO of the AV company should face prison time for their role in the death of the pedestrian. There was no significant main effect of company size or the interaction effect between oversight condition and company size ($ps > .34$). The main effect was significant, $\chi^2(3) = 10.94$, $p = .012$. When the *AI was not engaged*, participants were least likely to consider the CEO deserving of prison time (27%), but they thought s/he was considered to be significantly more deserving of prison when *oversight was not required* (49%, $p = .004$) or when *oversight was not mentioned* (45%, $p = .019$), with the *oversight required* condition falling in between, though without being reliably different (33%, $p = .36$). However, when the AI was engaged, participants were less likely to say that the CEO should be imprisoned when *oversight was required* compared to when *oversight was not required*, $p = .044$. Similar to the driver, company CEO attributions of criminal responsibility were positively correlated with general company responsibility ($r = 0.40$, $p < .001$).

Discussion

As in previous research (Bennett et al., 2020; Pöllänen et al., 2020), the company (which manufactured and programmed the vehicle) was held more responsible for the accident when control over the vehicle shifted away from the human driver. However, the human driver was always blamed to a non-trivial extent. Both when the vehicle was operating

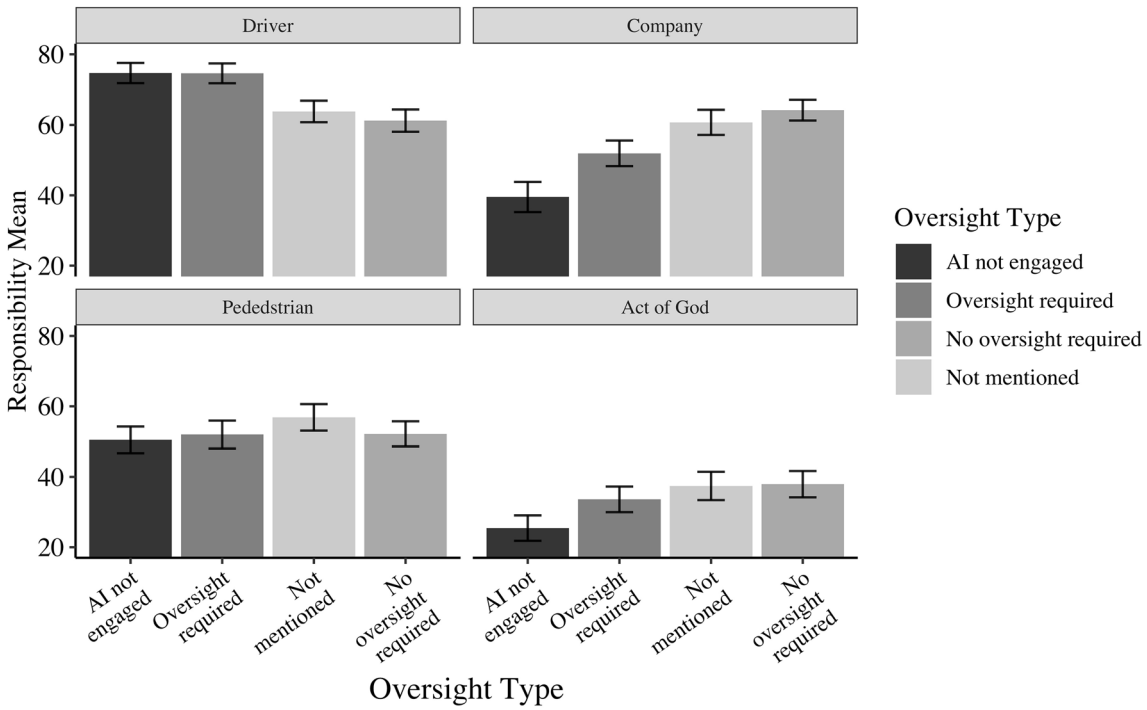


Fig. 1 Responsibility attributions to driver, company, pedestrian and act of God (Study 1)

autonomously and it was explicit that no oversight was required, and when no message concerning required oversight was conveyed (and people assumed that the vehicle was fully autonomous), drivers were still considered as responsible for the crash as was the company. At the same time, drivers were considered most responsible for the accident when they had full manual control over the vehicle (and the AI was not engaged), and even when they merely had the obligation to maintain oversight at all times. This finding is consistent with Pöllänen et al. (2020), who found that the total amount of blame leveraged against the manufacture or the vehicle increased for a partially autonomous vehicle compared to a traditional one, while the blame attributed to the human driver did *not* decrease.

At the same time, as in Pöllänen et al. (2020), the responsibility attributed to the company increased significantly when the vehicle was fully automatic and did not require any oversight. This effect seemed to be accompanied by a significantly lower inclination to hold the human driver responsible, even when the mean of the rating remained in the upper half of the scale. Whereas responsibility attributions to the pedestrian varied little across conditions, attributions to an act of God were elevated whenever the vehicle was operating autonomously, regardless of whether oversight was required or not. That is, only when the driver had full manual control did participants seem to exclude the possibility of an uncontrollable and unforeseeable cause. This finding suggests that participants seemed to have a considerable tolerance for error when AI was involved, but a much lower threshold for human drivers.

Company size matters primarily, though not exclusively, when the driver was operating an AV without the AI having been engaged. In all instances, the driver was blamed the most, regardless of company size. However, larger companies were clearly blamed much less than smaller companies. No other differences as a function of company size emerged.

Study 2

The goal of our second study was to replicate our Study 1 findings and assess their robustness. We asked participants not only to rate the responsibility of each party for the accident, but also to indicate the level of blame. In Study 2 we also altered the language used to manipulate company size, in case our Study 1 effect suffered from too weak a manipulation, and we used different language when asking participants to attribute blame to an act of God. Lastly, we explored various moderator variables including political orientation, gender, and age, as well as perceptions of the safety of AVs in general, all of which were assessed before participants encountered the accident vignette. However, because none of these moderator variables resulted in significant findings, they are not reported on further.

Method

Participants

We recruited 268 participants from MTurk and compensated them $0.75. Our final sample was 260 participants after removing seven participants with duplicate IP addresses and one participant who took fewer than 60 s to complete the survey (completion time $M = 415$ s, $Md = 316$ s). The sample self-identified as 70% men and included 68.5% Caucasians/Whites, 13.1% African Americans/Blacks, 10.0% Latin/Hispanic, 4.6% Asians. The mean age was 36.5 ($SD = 11.1$; range = 18–70).

Design

Study 2 employed the same design as Study 1, that is a 4 (oversight: required vs. not required vs. not mentioned vs. AI not engaged) × 2 (company size: small vs. large) between-groups factorial design.

Procedure

All procedures were identical to Study 1 except for the following amendments. First, Study 2 used a separate rating of blame following the responsibility ratings. Blame ratings were highly correlated with responsibility ratings, $r > .80$. Because we did not discern different patterns of results, we only report results for responsibility (see Appendix Part B for blame rating results). Second, in case in Study 1 positioning the slider at the mid-point scale (50 on a 0 to 100 scale) had inflated responsibility attribution, in Study 2 the starting position on the slider scales used for responsibility attributions was changed to 0 (see Table 1 for responsibility means for Study 1 and Study 2). That is, all participants needed to move the slider in order to assign any level of responsibility. Third, we added a slight modification of our oversight manipulation. Whereas in Study 1 we had pointed out that the vehicle that was fully self-driving without requiring human oversight was still "in the final development" (a detail not attributed to the vehicle in the other experimental conditions), we omitted any mention of the developmental stage in Study 2. Fourth, in light of the surprising main effect of company size, Study 2 altered the manipulation for company size to re-examine the nature of the effect. Lastly, whereas in Study 1 we interpreted our attribution to "no one in particular" as an act of God, Study 2 mentioned this term explicitly when participants made attribution judgments.[8] See Appendix Part C for a summary of all method-related variations between Study 1 and Study 2.

[8] The exact wording of the revised item was "No one is to blame for the accident, (i.e. this was an 'act of God')."

Results

Responsibility

As in Study 1, we submitted data to a linear mixed effects model using the *lme4* package in R (Bates et al., 2015). Unlike before, there was a main effect of oversight condition, $F(3, 252) = 4.15$, $p = .03$. Participants attributed more responsibility across all parties when *oversight was not mentioned* ($M = 53.08$, $SD = 33.13$, $p = .01$) or *oversight was required* ($M = 52.19$, $SD = 29.77$, $p = .02$) than when the *AI was not engaged* ($M = 44.47$, $SD = 33.39$). The condition in which *oversight was not mentioned* did fall in-between the others without being reliably different from *oversight was not required* and *oversight required*, $ps > .53$. We did not replicate the main effect of company size from Study 1 ($p = .21$). However, we did replicate a significant main effect of which party participants held more responsible $F(3, 756) = 46.52$, $p < .001$. Consistent with previous findings by Awad et al. (2020), participants attributed higher levels of responsibility to the driver in comparison to the company, pedestrian, and act of God, all $ps < .001$ (see Table 1).[9] Participants also attributed higher levels of responsibility to the company and pedestrian than to an act of God ($ps < .0001$). Critically, these two main effects were qualified by an interaction, $F(9, 756) = 4.43$, $p < 0.001$ (see Fig. 2).

Most Held Responsible Replicating Study 1, the driver was always held most responsible when the *AI was not engaged*, $p < .001$. Inconsistent with our prediction, the driver, the company, and the pedestrian were held equally responsible when *oversight was required* ($ps > .16$), though all were considered more responsible than an act of God ($p < .001$). These results suggest participants still held the company accountable even though the driver was supposed to oversee the vehicle.

When *oversight was not required*, the driver, the company, and the pedestrian were also held equally responsible ($ps > .08$), though again each were more responsible than an act of God ($p < .001$). This finding only partially replicated Study 1, which found higher responsibility attributions for the driver and the company than the pedestrian. Finally, there was no reliable difference in responsibility attributions between the company and both the driver and pedestrian when *oversight was not mentioned* ($p > .22$), though drivers in this condition were held more responsible than pedestrians ($p = .03$) and an act of God was rated the least responsible ($ps < .001$). Results mostly replicated Study 1, with the exception that in Study 2 the driver was held more responsible than the pedestrian when *oversight was not mentioned*, whereas in Study 1 there was no difference. (For ease of comprehension, the same Study 2 data as in Table 1 and Fig. 2 are also displayed in Appendix Part D, allowing for a direct comparison between attributions of responsibility to different sources of agency.)

Driver Responsibility Participants attributed more responsibility to the driver when the *AI was not engaged* ($M = 65.49$, $SD = 26.17$), $p = .003$, when *oversight was not mentioned* ($M = 63.58$, $SD = 26.24$), $p = .008$, and when *oversight was required* ($M = 60.27$, $SD = 25.13$), $p = .04$, compared to when *oversight was not required* ($M = 49.75$, $SD = 25.40$). This result suggests that the driver is not held as responsible when the company indicated that the driver does not need to oversee the AV. Replicating Study 1, participants considered the driver to be more responsible when *oversight was required* and when *AI was not engaged* than when *oversight was not required*. However, we did not replicate Study 1 findings that drivers received more responsibility in the *oversight required* and the *AI not engaged* conditions compared to when *oversight was not mentioned* ($ps > 0.55$). Additionally, Study 1 participants attributed similar responsibility to the driver in the *oversight not required* and *oversight not mentioned* conditions, whereas in Study 2 participants attributed more responsibility in the oversight required conditions.

Company Responsibility The company was held more responsible when *oversight was required* ($M = 59.34$, $SD = 31.46$), *oversight was not mentioned* ($M = 58.08$, SD = 29.81), or when *oversight was not required* ($M = 56.60$, $SD = 29.13$) compared to when the *AI was not engaged* ($M = 34.11$, $SD = 31.46$) $ps < .0001$. These findings suggest companies are held responsible in all conditions where the AI is on, regardless of whether they require the driver to oversee the AV. These findings replicate Study 1. However, this study did not replicate our earlier finding that companies are held more responsible when *oversight was not required* compared to when *oversight was required*. This finding suggests that companies are not relieved of responsibility when they require the driver to maintain oversight over the AV at all times.

Act of God and Pedestrian Unlike Study 1, participants did not differ in considering the accident an act of God by levels of oversight. However, we replicated findings from Study 1 that there were no differences in pedestrian responsibility across levels of oversight.

Criminal Liability

Driver As indicated in Table 1, drivers were not held as responsible in Study 2 as in Study 1, possibly due to the difference in slider position. Still, Study 2 found that 58% of our sample believed that Kate should be held criminally

[9] Overall, in spite of the slightly different online questionnaire design and the varying order of measures, means for the responsibility ratings were indistinguishable between Study 1 and Study 2. The only exception were the driver responsibility ratings, which were significantly higher in Study 1 than in Study 2 ($\Delta M = 8.47$, 95%CI [4.11, 12.83]).

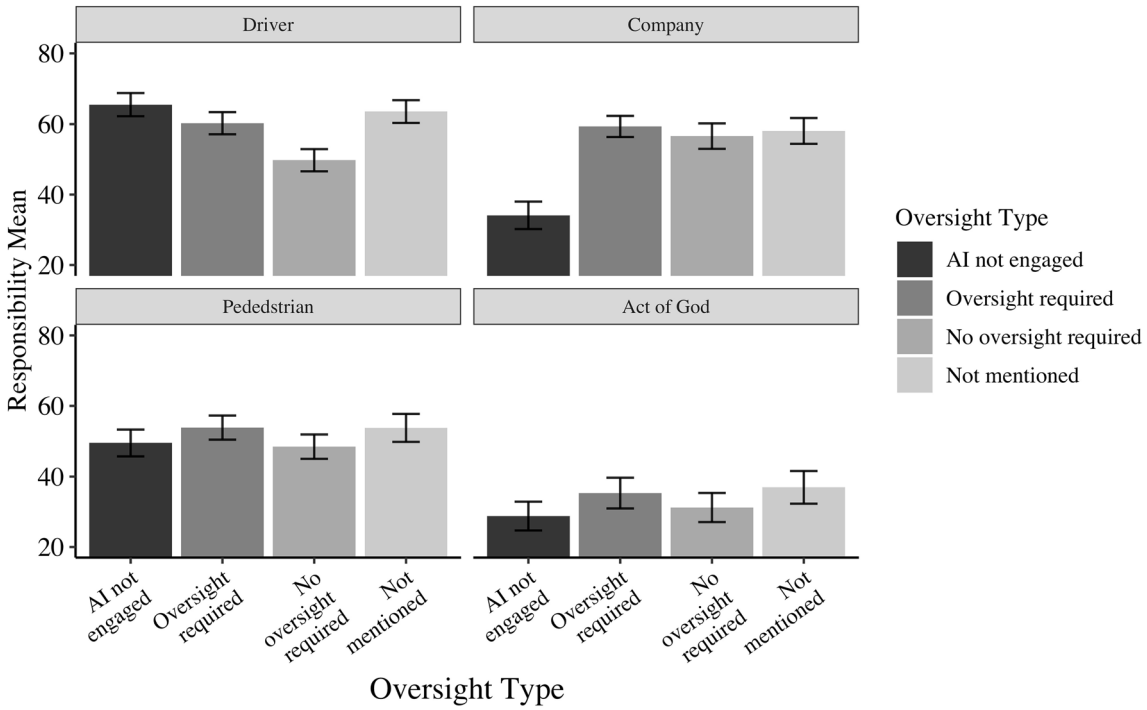


Fig. 2 Responsibility attributions to driver, company, pedestrian and actor of God (Study 2)

responsible, and these effects were marginally dependent upon the oversight condition, $\chi^2(3) = 6.64$, $p = .08$. Within the oversight condition, the effect from Study 1 was somewhat replicated in regard to the *AI not engaged* condition. Participants in this condition (71%) were significantly more likely than those in the *no oversight required* condition (49%) to believe the driver should be held criminally responsible, $p = .01$, and slightly more likely to believe the driver should be held criminally responsible than the *oversight not mentioned* category (56%, $p = .08$). The difference did not hold for the *oversight required condition* (58%, $p = .13$). These findings suggest that the act of having any level of AI autonomy engaged relieves the driver of some criminal responsibility in the minds of research participants. Driver criminal responsibility was positively correlated with general driver responsibility ($r = 0.34$, $p < .001$), suggesting participants' responsibility attributions are related to their beliefs regarding criminal responsibility.

Company CEO In Study 2, 48% of participants believed the company CEO should be held criminally responsible for the accident. The oversight manipulation significantly affected these responsibility attributions, $\chi^2(3) = 9.44$, $p = .02$. Replicating Study 1, participants were less likely to believe the CEO should be held criminally responsible when the *AI was not engaged* (32%) than when *oversight was not mentioned* (56%, $p = .007$) or when *oversight was required* (55%, $p = .01$). Similar to Study 1, participants differed only marginally between the *AI not engaged* and *oversight not required* conditions in whether they believed the CEO should go to prison. Those in the *oversight required* condition were slightly more likely than those in the *oversight not required* condition (48%) to believe the CEO should be held criminally responsible, $p = .07$. This was contrary to expectations, as a company that had explicitly delegated some of the responsibility to the driver should be held less responsible than a company that accepted all responsibility from the start. There were no other significant differences, $ps > 0.33$. Company CEO criminal responsibility was also positively correlated with general company responsibility ($r = 0.34$, $p < .001$).

Discussion

Our second investigation reproduced most of the patterns from Study 1. Drivers, on average, were more responsible for the accident than companies. Even with a slightly different manipulation, company size, again, was inconsequential. Still, some noteworthy differences emerged.

As in Study 1, the company was more responsible when any level of automation was engaged. Yet, recall that in Study 1 participants held the company most responsible when no oversight was required compared to when the driver was obligated to pay attention to the road. In our second study, however, participants did not seem to distinguish between levels of oversight in terms of company responsibility. It is difficult to interpret this apparent discrepancy. One could speculate that

methodological details between Study 1 and Study 2 may have contributed to this difference. In both studies, participants indicated their responsibility attributions with a slider tool, which by default was set to the middle of the scale (50) in Study 1, but to 0 in Study 2. Pertinent research by Liu and Conrad (2019) shows that a slider's starting position does not affect participants' response, except when a 100-point scale is used. This is, however, precisely the scale of choice in the present research. Liu and Conrad (2019) demonstrated that on such 100-point scales, participants are likely to accept the default position as their response.

However, few participants in our studies accepted the default slider position. Especially in Study 2, with the slider set to 0 from the beginning, all participants had to move the slider in order to assign any level of responsibility. In Study 1, across all four sliders only 1.3–5.5% of all responses reflected the default response of 50 on a 100-point scale, arguably a negligible number. In Study 2, participants were only likely to leave the slider at 0 when this was a plausible response. Participants blamed the company the least when the driver had turned off the AI, with 31% of all responses being 0, or when 32% of all participants rejected the possibility that this was an act of God. Corresponding percentages of "0" responses from Study 1 were 37% and 42%. Hence, the default slider position is extremely unlikely to have shaped the difference in findings between Study 1 and 2. At this stage, we do not have enough information to determine which result deserves greater credence; thus, more research is necessary.

Still, the default slider position may have had some effect because the overall level of driver responsibility was substantially higher in Study 1 than Study 2, consistent with the default providing a higher anchor in Study 1. At the same time, there did not appear substantive variation in the studies across pedestrian responsibility and acts of God.

Other differences between studies included the oversight not mentioned condition, such that in Study 2 drivers were held responsible at a level similar to the AI not engaged group, whereas in Study 1 driver responsibility was comparable to driver responsibility in the oversight not required condition, though results for act of God and pedestrian blame were consistent across investigations. Findings for criminal responsibility for the driver and company CEO largely mirrored those of Study 1.

General Discussion

The present findings contribute to a growing literature on responsibility attributions concerning automated driving. First, human drivers were disproportionately blamed for the accident. The responsibility attributed to human drivers was never significantly lower than the responsibility attributed to the company, even when the AI was operating the vehicle; that is, at minimum drivers were blamed at roughly the same level as the company even when the driver was technically relieved of some responsibility during the operation of the vehicle. Conversely, companies *were* held less responsible when the AI was turned off by the driver, and blamed more when the AI was turned on. Whereas the precise level of blame allocated varied based on whether the human driver retained any oversight duty, overall the pattern highlights a stark asymmetry: Humans are blamed more than machines in the context of a deadly accident. Our research adds to the findings by Awad et al. (2020), Bennett et al. (2020) and Pöllänen et al. (2020), all of whom have elucidated different aspects of this asymmetry.

Why does this asymmetry occur? In the extant research on attribution and causal explanation, it is well established that human agency is often considered a sufficient cause for an event to have occurred (e.g., Hilton et al., 2010; Lagnado & Channon, 2008). Part of this prominence in causal reasoning might be counterfactual thinking: It is comparatively easy for people to imagine that a human driver "could have done something differently" than to imagine different behavior from a machine, for which it must be assumed does not have any choice in the sense that its responses are predetermined by algorithms. By the same token, especially in the context of driving, observers may feel that human drivers are or should be aware of any alternate courses of action, and be able to anticipate the consequences of different actions (see e.g., Hilton et al., 2016, and Lagnado & Channon, 2008 for the critical role of foreseeability in judgments of responsibility). This is not necessarily the case for machines. Put differently, at least in the conception of lay people, machines do not "decide"; machines follow a preprogrammed sequence of if's and then's, but only in exceptional circumstances do they encounter a situation in which the machine's response was not determined in advance.

However, even if people imagine AI as being reflective of human agency, the perceived psychological distance of the agent to the accident may steer observers in the direction of putting responsibility on human drivers. Human drivers are present in the situation and located behind the wheel. By comparison, any agency imbued in the AI of the vehicle may ultimately reside with the programmer and the manufacturer, who determined how the vehicle would respond. This implies that the AI in the AV is only a proxy for the actual responsible agents. With these agents being removed from the immediate proximity of the accident, it might not be surprising these agents/"the company" is blamed less as these agents are perceived as less instrumental in the accident. Indeed, Li et al. (2016) found that, although the vehicle itself was perceived as more responsible when it was fully autonomous than manual, fully autonomous vehicles were not held as responsible as human drivers of manually driven vehicles. Though speculative, the results from our research (and Li et al.' (2016)

findings) could suggest people might view the AI programmer and manufacturer as having less direct control over any given outcome, and therefore people attribute less responsibility toward the AV programmer or manufacturer. In short, people tend to be considered responsible because they are assumed to be able to foresee the consequences of their actions (or the actions they allow the AV to engage in on their behalf), and because they are directly involved in the situation, implying knowledge and control.

A related consideration is that observers might be skeptical as to whether a "no need to supervise" assurance absolves the driver of moral responsibility. That is, if an AV manufacturer tells vehicle owners that they do not have to pay attention, vehicle operators might believe that this absolves them from any moral responsibility to take control of the vehicle. However, an observer who is not involved in the immediate operation of the vehicle may not absolve the vehicle operator from responsibility. Rather, human agency might feature prominently in the perspective of the observer given that the vehicle owner has still instructed the AV to arrive at a specific destination, therefore initiating and authorizing the AV's actions. That is, counterfactually the driver is necessary for this specific accident to have occurred in the particular time and place. With people typically rendering judgments of responsibility based on salient counterfactuals, we surmise that drivers will still be held accountable to the extent that the operation of the AV can be construed as being somewhat under the driver's control and as an extension of their intent. Practically speaking, even when there should be fully autonomous vehicles in the future, human drivers will likely be deemed at least partially responsible for any accidents that occur.

Vehicle manufacturers are cognizant of the fact that drivers of an AV will become accustomed to the vehicle taking care of all necessary operations. More generally, this leads to what Triberti et al. (2020) have deemed "role ambiguity": Even though humans are technically in charge, practically their intervention is not required, allowing them to direct their attention elsewhere. Because of this tendency, some manufacturers have installed devices to ensure that human drivers do pay attention (e.g., Thompson, 2016). Independent of whatever verbal instructions may have been conveyed to the human driver, the actual sensory experience of the situation communicates to the person that their attention is not needed, rendering it understandable that they are not ready to intervene. Because our vignette format did not convey this sensory experience, we must assume that our research participants may overestimate the level of attention that is plausible for a human driver in this situation to maintain. In short, even though drivers were told nominally to oversee the vehicle and intervene at any time, doing so might not have been realistic. If participants had any experience overseeing an AV, they might have more understanding of the lack of plausibility in overseeing the vehicle at all times, and therefore may have been more lenient in their responsibility attributions.

The tendency to focus on human agency in the attribution of responsibility has important implications. However rarely that situation may arise, people may anticipate being blamed for accidents in which they have no immediate control, and this might stifle their willingness to adopt AVs. Put differently, people might be resistant to life-saving technology if they are formally responsible for the mistakes made by a non-human agent. Elish (2019) refers to a situation when responsibility is misattributed to a human actor with limited control in an automated system as a "moral crumple zone." The notion of a moral crumple zone posits that people will focus more on the immediate human driver when determining the actor responsible for an accident, and therefore less on indirect human actors (e.g., programmers and manufacturers). This may have implications for delaying AV innovation in two central ways. First, a tendency to blame the human driver can stifle insight into what caused the accident. As with any complex system there are likely multiple causes to an accident and holding only one actor accountable is an oversimplification. Second, the decreased blame placed on machines for AV accidents could result in a lack of public pressure on governments to propose new regulations, or simply consumer pressure that forces manufacturers to improve their design (Awad et al., 2020). Therefore, the lack of sufficient insight into what caused an accident and a lack of pressure on governments to propose safety regulations can cause delays in improving AV systems.

An important observation in our research was that the company is still blamed to a considerable extent even when requiring a human driver to oversee the vehicle. Offhand, one would expect that this finding should have an exculpatory effect for the company in that it might be blamed at roughly the same level as when the AI is not engaged. However, participants clearly assumed a degree of shared responsibility between human driver and AI, notwithstanding any verbal instructions.

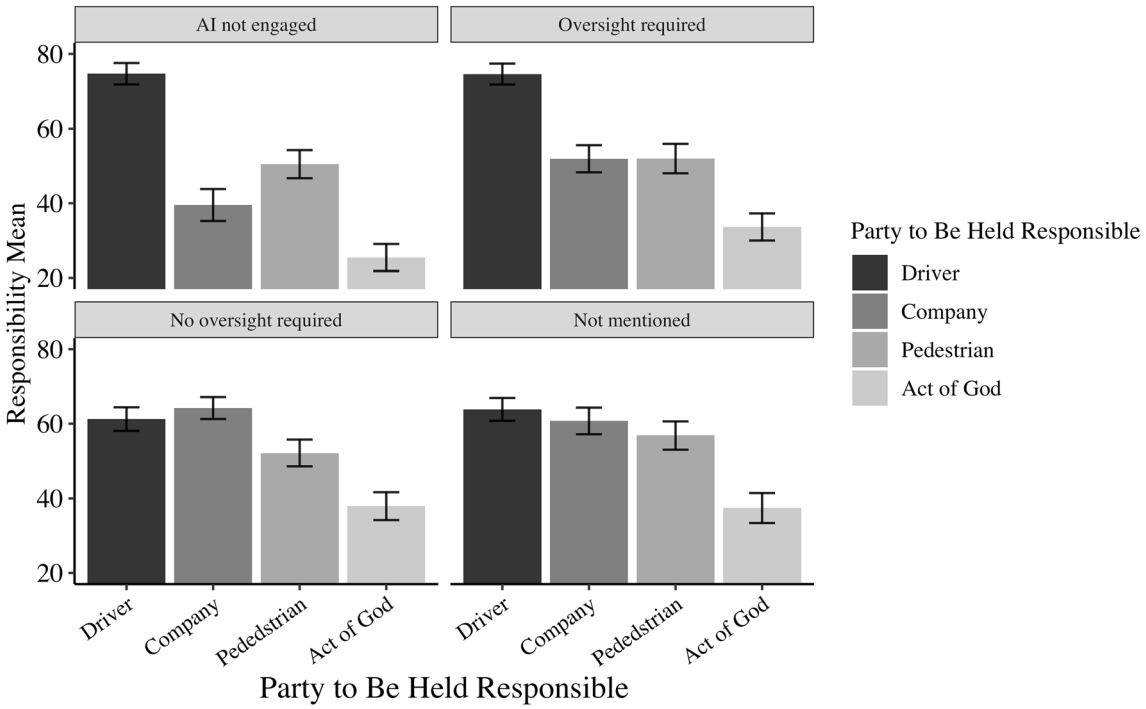
Another noteworthy aspect of our findings concerns criminal liability. Even when the AI was turned off, and regardless of whether participants recommended that the driver be held criminally responsible, between 27%–32% of participants wanted the company's CEO to risk serving prison time. When the AI was turned on, and even when the company required the driver to exercise oversight, still between 38%–55% of participants considered the company criminally responsible. Whereas such patterns corresponded to expected differences between experimental conditions, there is a readiness to hold AV companies legally accountable regardless of what took place. This might create a legal and economic risk for companies, which may delay the introduction of AVs as a life-saving technology. To the extent that the introduction of AV technology to reduce traffic risks is a priority, institutional solutions are required.

One possible solution is that each vehicle (through its company or manufacturer) should also carry insurance, just as it is obligatory for human drivers to be insured. This is especially relevant in situations when a human feels that they are being blamed disproportionately, and see the actual responsibility with the vehicle (e.g., when a person is unable to comply with the verbal instructions, see above). This, however, might increase the cost of operating AVs, and thus might be a deterrent for large segments of the population to adopt AVs. Although the present research cannot provide a comprehensive legal analysis of liabilities, it nevertheless points to a considerable degree of complication which must be legally regulated before companies (manufacturers) and human drivers are willing to treat AVs as a part of everyday life in the way that is currently the case for traditional automobiles.

Appendix

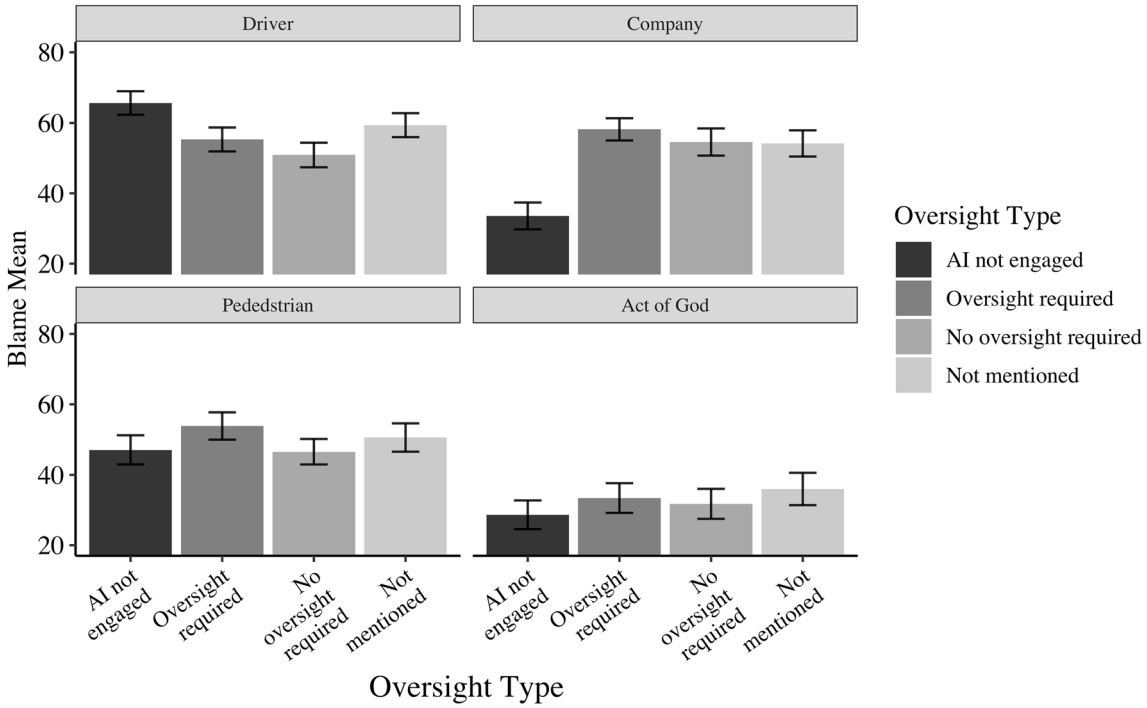
Part A

Responsibility attributions Study 1, allowing for a direction comparison of attributions to driver, company, pedestrian and act of God, by experimental condition



Part B

Blame attributions to driver, company, pedestrian and act of God (Study 2)



Study 2

Blame

The multilevel model results indicated a significant main effect of the party to be blamed $F(3, 756) = 39.71$, $p < 0.001$, such that participants attributed higher levels of blame to the driver ($M = 57.82$, $SD = 27.79$) compared to the company ($M = 50.12$, $SD = 30.86$), pedestrian ($M = 49.50$, $SD = 31.34$), and act of God ($M = 32.47$, $SD = 34.58$; $ps \leq .001$). Participants also attributed higher levels of blame to the company and pedestrian than an act of God ($ps < .0001$). Critically, this effect was qualified by a significant interaction between oversight and the party to be blamed $F(9,756) = 4.19$, $p < .001$ (see Fig. 3). Pairwise analyses indicated attributions of blame to the driver and company differed by levels of oversight. Participants attributed more blame to the driver when the AI was not engaged ($M = 65.63$, $SD = 26.57$, $p = .007$) than when oversight was not required ($M = 50.89$, $SD = 27.20$).

Similar to attributions of responsibility, participants attributed more blame to the company when oversight was required ($M = 58.19$, $SD = 25.33$), was not mentioned ($M = 54.18$, $SD = 30.22$), or when no oversight was required ($M = 54.58$, $SD = 31.11$) compared to when the AI was not engaged ($M = 33.57$, $SD = 30.76$, $ps < .0002$). These findings suggest companies are blamed in all conditions where the AV is turned on, regardless of whether they require the driver to oversee the AV. Attributions of blame to the pedestrian or an act of God did not differ by levels of oversight.

Part C

Table 2 Differences in Method and Materials between Study 1 and Study 2

Material changed	Study 1	Study 2	Explanation for change
Blame measure	Not included	Included	To measure potential differences between blame and responsibility
Company size manipulation (small vs large)	"*Small and upcoming*" vs. "*large and well established*"	"*Small automotive startup company with 300 employees in the United States*" vs. "*large multinational automotive company with 46,000 employees across five continents*"	To increase the effectiveness of the company size manipulation
Placement of potential moderators	After vignette	Before vignette	To account for moderator responses varying by condition in Study 1
Default Slider position	Mid-point (50 on a 0 to 100 scale)	0 (0 on a 0 to 100 scale)	In case the slider position at the mid-point inflated attributions of responsibility
'Act of God' attributions	"*No one is responsible for the accident; bad things just happen sometimes*"	"*No one is to blame for the accident, (*i.e. *this was an 'act of God')*"	To explicitly use the legal term "act of God"
Oversight manipulation (required vs not required)	"*Is still in the development phase and needs consistent human oversight*" vs. "*is in the final phase of development and does not need consistent human oversight*"	"*Needs consistent human oversight, with a driver paying attention and being able to intervene at all times*" vs. "*does not need consistent human oversight, with drivers not being required to pay attention or having to intervene*"	To explicitly define that human oversight entails paying attention and having the ability to intervene

Part D

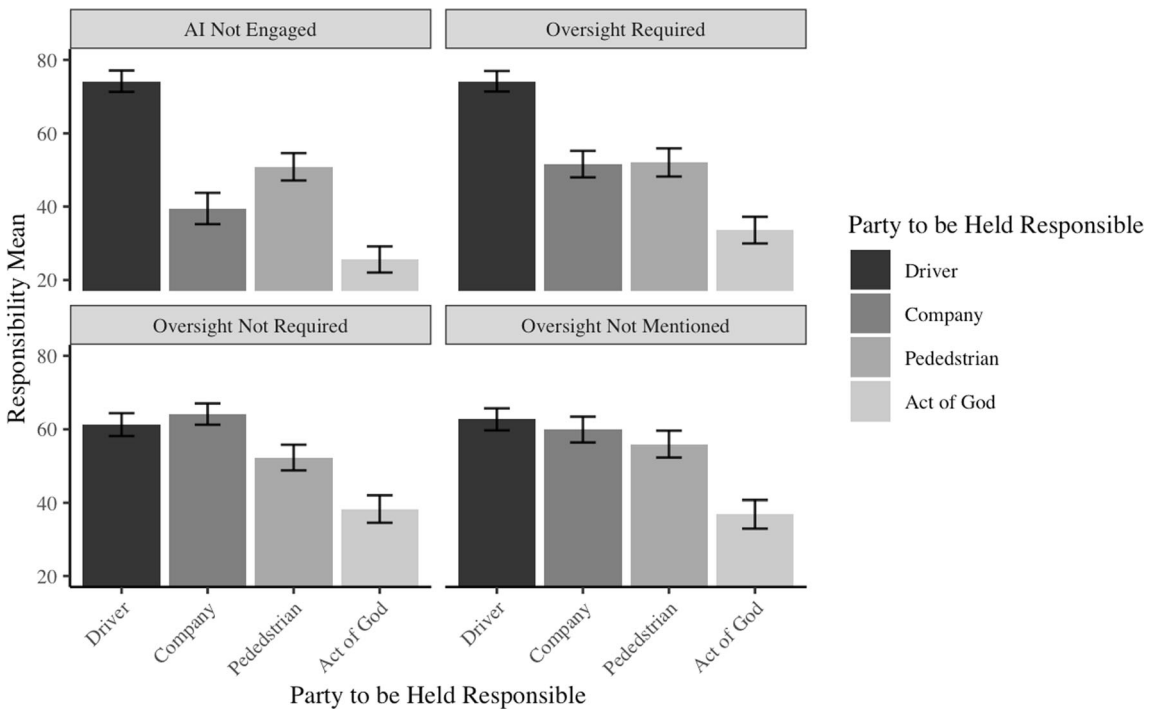


Fig. 3 Responsibility attributions Study 2 allowing for a direct comparisons of attributions to driver, company, pedestrian and act of God, by experimental condition

Acknowledgments This work was supported by the Nevada Center for Applied Research, Director Carlos Cardillo, Ph.D., through funding from the Intelligent Mobility program of the Regional Transportation Commission of Washoe County, Nevada. This research was also supported by a small grant from the National Judicial College, Reno, NV, President the Honorable Benes Aldana. We gratefully acknowledge Tyler Livingston for his comments on an earlier draft of this manuscript.

Data Availability The data for the two studies reported in this paper are available at the Open Science Foundation: https://osf.io/x4wp2/?view_only=bd7d059563804f43b35882b3d64b9dd4

Declarations

Ethics Approval The research reported in this article was determined to be exempt from review by the Institutional Review Board of the University of Nevada, Reno, Office of Research Integrity, IRBNET case number 1578869, March 27, 2020.

Conflict of Interest None of the authors has any conflict of interest.

References

Alicke, M. D. (2000). Culpable control and the psychology of blame. *Psychological Bulletin, 126*(4), 556–574.

Arizona's Executive Order 9, 2015, or). https://apps.azdot.gov/files/sitefinity-files/Executive-Order-2015-09.pdf

Awad, E., Levine, S., Kleiman-Weiner, M., Dsouza, S., Tenenbaum, J. B., Shariff, A., Bonnefon, J. F., & Rahwan, I. (2020). Drivers are blamed more than their automated cars when both make mistakes. *Nature Human Behaviour, 4*(2), 134–143.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.

BBC News. (2019, November 6). Uber in fatal crash had safety flaws say US investigators. . Retrieved from https://www.bbc.com/news/business-50312340

Bennett, J. M., Challinor, K. L., Modesto, O., & Prabhakharan, P. (2020). Attribution of blame of crash causation across varying levels of vehicle automation. *Safety Science, 132*, 1–12. https://doi.org/10.1016/j.ssci.2020.104968.

Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science, 352*(6293), 1573–1576. https://doi.org/10.1126/science.aaf2654.

Carroll, J. S., Perkowitz, W. T., Lurigio, A. J., & Weaver, F. M. (1987). Sentencing goals, causal attributions, ideology, and personality. *Journal of Personality and Social Psychology, 52*, 107–118. https://doi.org/10.1037/0022-3514.52.1.107.

De Freitas, J., Anthony, S. E., Censi, A., & Alvarez, G. A. (2020). Doubting driverless dilemmas. *Perspectives on Psychological Science, 15*(5), 1284–1288. https://doi.org/10.1177/1745691620922201.

Elish, M. C. (2019). Moral crumple zones: Cautionary tales in human-robot interaction. *Engaging Science, Technology, and Society, 5*, 40–60.

Elish, M. C. (2020, June 15). Who is responsible when autonomous systems fail? Centre for international governance innovation. Retrieved from https://www.cigionline.org/articles/who-responsible-when-autonomous-systems-fail

Foot, P. (1967). The problem of abortion and the doctrine of double effect. *Oxford Review, 5*, 5–15.

Gonzales, R. (2019, November 7). Feds say self-driving Uber SUV did not recognize jaywalking pedestrian in fatal crash. NPR. Retrieved from https://www.npr.org/2019/11/07/777438412/feds-say-self-driving-uber-suv-did-not-recognize-jaywalking-pedestrian-in-fatal-.

Hans, V. P. (2000). *Business on trial*. Yale University Press.

Heinze, J., Uhlmann, E. L., & Diermeier, D. (2014). Unlikely allies: Credibility transfer during a corporate crisis. *Journal of Applied Social Psychology, 44*, 392–397. https://doi.org/10.1111/jasp.12227.

Hilton, D. J., McClure, J., & Sutton, R. M. (2010). Selecting explanations from causal chains: Do statistical principles explain preferences for voluntary causes? *European Journal of Social Psychology, 40*(3), 383–400.

Hilton, D. J., McClure, J., & Moir, B. (2016). Acting knowingly: Effects of the agent's awareness of an opportunity on causal attributions. *Thinking & Reasoning, 22*(4), 461–494.

Lagnado, D. A., & Channon, S. (2008). Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition, 108*(3), 754–770. https://doi.org/10.1016/j.cognition.2008.06.009.

Li, J., Zhao, X., Cho, M., Ju, W., & Malle, B. (2016). From trolley to autonomous vehicle: Perceptions of responsibility and moral norms in traffic incidents with self-driving cars. Paper presented at the Society of Automotive Engineers World Congress, Detroit, MI. https://doi.org/10.4271/2016-01-0164.

Liu, M., & Conrad, F. G. (2019). Where should I start? On default values for slider questions in web surveys. *Social Science Computer Review, 37*(2), 248–269.

MacCoun, R. J. (1996). Differential treatment of corporate defendants by juries: An examination of the" deep-pockets" hypothesis. *Law and Society Review, 30*, 121–161.

McCausland, P. (2019, November 9). Self-driving Uber car that hit and killed woman did not recognize that pedestrians jaywalk. NBC News. Retrieved from https://www.nbcnews.com/tech/tech-news/self-driving-uber-car-hit-killed-woman-did-not-recognize-n1079281

McManus, R. M., & Rutchick, A. M. (2019). Autonomous vehicles and the attribution of moral responsibility. *Social Psychological and Personality Science, 10*(3), 345–352.

National Highway Traffic Safety Administration (2015, February). *Critical reasons for crashes investigated in the national motor vehicle crash causation survey*. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115

National Safety Council. (2020). *Motor vehicle deaths estimated to have dropped 2% in 2019*. https://www.nsc.org/road-safety/safety-topics/fatality-estimates

National Transport Commission (2020). Automated vehicle program approach. https://www.ntc.gov.au/sites/default/files/assets/files/Automated%20vehicle%20approach.pdf

Nevada Revised Statue 482A. https://www.leg.state.nv.us/NRS/NRS-482A.html

Pöllänen, E., Read, G. J. M., Lane, B. R., Thompson, J., & Salmon, P. M. (2020). Who is to blame for crashes involving autonomous vehicles? Exploring blame attribution across the road transport system. *Ergonomics, 63*(5), 525–537. https://doi.org/10.1080/00140139.2020.1744064.

Randazzo, R. (2019, March 17). Uber crash death in Tempe: A closer look. AZ Central. Retrieved from https://www.azcentral.com/story/news/local/tempe/2019/03/17/uber-crash-death-who-blame-tempe-arizona-rafaela-vasquez-elaine-herzberg/3157481002/

Schlenker, B. R., Chambers, J. R., & Le, B. M. (2012). Conservatives are happier than liberals, but why? Political ideology, personality, and life satisfaction. *Journal of Research in Personality, 46*(2), 127–146. https://doi.org/10.1016/j.jrp.2011.12.009.

Shaver, K. G. (2012). The attribution of blame: Causality, responsibility, and blameworthiness. Springer Science & Business Media.

Thompson, C. (2016, September 9). GM is taking on Tesla's Autopilot with a self-driving system that tracks your eyes. . Retrieved from https://www.businessinsider.com/general-motors-super-cruise-eye-tracking-technology-2016-9

Triberti, S., Durosini, I., & Pravettoni, G. (2020). A "third wheel" effect in health decision making involving artificial entities: A psychological perspective. *Frontiers in Public Health, 8*, 117. https://doi.org/10.3389/fpubh.2020.00117.

Vidmar, N. (1999). The performance of the American civil jury: An empirical perspective. *Arizona Law Review, 40*, 849–899.

Wamsley, L. (2019, March 6). Uber not criminally liable in death of woman hit by self-driving car, prosecutor says. NPR. Retrieved from https://www.npr.org/2019/03/06/700801945/uber-not-criminally-liable-in-death-of-woman-hit-by-self-driving-car-says-prosec

Weiner, B. (1995). Judgments of responsibility: A foundation for a theory of social conduct. Guilford.