



Moral judgment, decision times and emotional salience of a new developed set of sacrificial manual driving dilemmas

Giovanni Bruno¹ · Michela Sarlo² · Lorella Lotto³ · Nicola Cellini^{1,4,5,6} · Simone Cutini^{3,5} · Andrea Spoto¹

Accepted: 11 November 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

The growing interest in the subject of moral judgment in driver and autonomous vehicle behavior highlights the importance of investigating the suitability of sacrificial dilemmas as experimental tools in the context of traffic psychology. To this aim a set of validated sacrificial trolley problems and a new set of trolley-like driving dilemmas were compared through an online survey experiment, providing normative values for rates of participants' choices; decision times; evaluation of emotional valence and arousal experienced during the decision process; and ratings of the moral acceptability. Results showed that while both sets of dilemmas led to a more frequent selection of utilitarian outcomes, the driving-type dilemmas seemed to enhance faster decisions mainly based on the utilitarian moral code. No further differences were observed between the two sets, confirming the reliability of the moral dilemma tool in the investigation of moral driving behaviors. We suggest that as moral judgments and behaviors become more lifelike, the individual's moral inclination emerge more automatically and effectively. This new driving-type dilemma set may help researchers who work in traffic psychology and moral decision-making to approach the complex task of developing realistic moral scenarios more easily in the context of autonomous and nonautonomous transportation.

Keywords Utilitarianism · Moral decision-making · Moral Dilemma · Sacrificial dilemma · Driving behavior

Introduction

During the last twenty years, the enthusiasm in the traffic psychology research field has newly flourished, mainly fueled by the worldwide interest in the cognitive, social and ethical mechanisms behind driving behavior in vision of the

upcoming revolution of autonomous transportation (e.g., Baum, 2020; De Sio, 2017; Gkartzonikas & Gkritza, 2019; Haboucha et al., 2017; Hidalgo et al., 2021; Pigeon et al., 2021; Shariff et al., 2017).

Focusing on the ethical framework, several studies have investigated the beliefs that people have about how autonomous and human drivers should behave when facing complex decisions in which moral values and rules are involved (e.g., Bonnefon et al., 2016; Goodall, 2014; Meder et al., 2019; Sütfeld et al., 2019; Sütfeld et al., 2017). The definition of the "right" behavior is far from an easy task, considering how much individuals' and countries' ethical principles, social preferences, and traffic culture differ from each other (Awad et al., 2018; Graham et al., 2016, Özkan et al., 2006).

Regardless of the level of automation, a widely used experimental approach to investigating reasoning and moral judgment behind the wheel is the sacrificial moral dilemma (e.g. Martí-Vilar et al., 2021; Nasello et Al., 2021, Trémolière et al., 2017), defined as a conflict between two undesirable actions with unpleasant outcomes involving loss of human lives. The trolley and the footbridge problems are

✉ Giovanni Bruno
giovanni.bruno.4@phd.unipd.it

¹ Department of General Psychology, University of Padua, Via Venezia 8, 35131 Padua, Italy

² Department of Communication Sciences, Humanities and International Studies, University of Urbino Carlo Bo, Urbino, Italy

³ Department of Developmental Psychology and Socialization, University of Padua, Padua, Italy

⁴ Department of Biomedical Sciences, University of Padua, Padua, Italy

⁵ Padova Neuroscience Center, University of Padua, Padua, Italy

⁶ Human Inspired Technology Center, University of Padua, Padua, Italy

two paradigmatic examples (Foot, 1978; Thomson, 1985). In the first case, a trolley moving along tracks is on course to run down five track workers. The only way to save them is by pulling a lever that diverts the trolley onto another track where only one worker will die. In the footbridge version, instead, the five track workers can be saved only by pushing a large man off an overpass onto the main track: he will die, but his body will stop the trolley. The main aim of moral dilemmas such as these is to bring out two categories of different schools of moral thought: Kantian deontology and Benthamian utilitarianism. Deontology is a duty-based ethical theory that focuses on moral norms (i.e., categorical imperatives) that cannot be violated (e.g., “the ends never justify the means”; Kant, 1785). Utilitarianism, as the paradigm case of consequentialism, claims that an act is right if and only if it minimizes overall harm, denying that moral rightness depends on anything other than consequences (e.g., Bentham, 1781).

In this context, the exploration of the role of intention in moral decision-making is also thorough, on the basis of the doctrine of the double effect (DDE; Aquinas, 1952). According to this principle, voluntarily causing harm for a greater good (“instrumental dilemmas”) is not acceptable, albeit it is permissible as a foreseen—but unintended—side effect to save more people (“incidental dilemmas”). In line with the DDE, the footbridge dilemma is structured as an instrumental dilemma, in which the sacrifice of one person is the unethical means to save more people, and the trolley problem is structured as an incidental dilemma, in which diverting the course of the trolley towards the single worker is ethically permissible as a predictable—but undesired—consequence of achieving a greater goal (i.e., saving five people). The role of intention in moral judgments has been widely explored (e.g., Borg et al., 2006; Cushman et al., 2006), also through the manipulation of time pressure or cognitive interference (see Trémolière et al., 2017), and several studies have shown how participants were much more likely to cause harm unintentionally than intentionally (e.g., Borg et al., 2006; Cushman et al., 2006; Greene et al., 2001; Lotto et al., 2014; Moore et al., 2008).

A typical adaptation of this thought experiment to on-road scenarios consists of a driver—or an autonomous vehicle—facing a binary decision in a typical incidental dilemma: to maintain the trajectory and run over a number of pedestrians, killing all of them, or to swerve and kill a single unaware pedestrian or the driver/passenger (Bonneton et al., 2016). During the last decade, this hypothetical AV-type dilemma framework has been thorough in different experimental settings: virtual reality (e.g., Kallioinen et al., 2019; Sütfield et al., 2017, 2019), immersive driving simulations (e.g., Frison et al., 2016; Samuel et al., 2020), and importantly through text-based moral scenarios (e.g., Bonneton et al., 2016; Sütfield et al., 2019).

When the available actions were set for a dichotomous decision, the deontological option stood for a submissive action (e.g., maintaining trajectory), and the utilitarian one for a proactive action (e.g., swerving or breaking). In general, respondents strongly prefer vehicles programmed to maximize utility by saving the highest number of people (Awad et al., 2018). Among the several variations that have characterized on-road moral scenarios (e.g., manipulating the number or the ages of potential victims, their socioeconomic status, the road laws; Awad et al., 2018), the self-involvement risk factor seems a crucial application to the driving-type dilemmas (e.g., Bonneton et al., 2016; Sütfield et al., 2019). The seminal work by Bonneton et al. (2016) showed that drivers/passengers who approve utilitarian algorithms, when directly involved as potential victims, tend to prefer self-protective vehicles for themselves. Nonetheless, Bergmann and colleagues (2018) further explored the utilitarian approach when the decision maker’s life was at stake. Surprisingly, the participants acted more altruistically than expected when their life was at risk, protecting the pedestrians over themselves and proportionally to the number of potential victims. In the described experimental application, Bergmann and colleagues took the inspiration from Thomson’s dissertation on the absence of a self-sacrificial option in classic moral dilemmas (2008), which addition may cause a quarrel between utilitarian aims and instincts of self-preservation. Furthermore, in the exemplification of the moral agent’s three-option variant (killing five vs killing one vs killing himself), Thomson also utilized a typical on-road scenario, in which the car’s brakes fails and there was no chance to avoid a sacrificial solution. Evidence on this point are still controversial, and Thomson herself has opted to leave the floor to further investigations (i.e., Huebner & Hauser, 2011), underlining the complexity of interpreting an individual’s moral judgment when the self-preservation instinct collides with utilitarian collective demands.

In addition, emotional valence and arousal have an influence on driving behavior (Chan & Singhal, 2015), and the role that emotional processing plays in the development of the ‘right’ moral decision is well-known. Particularly, Greene’s dual-process theory underlined the systematic competition between emotional and cognitive processes in moral judgment: a higher endorsement of utilitarian options seems to be related to a lower emotional activation because a slow, controlled cognitive process can overwhelm a fast and automatic emotional response, when the perceived benefits of a moral decision exceed the costs (Greene et al., 2001, 2004, 2008). Specifically, some studies have focused on the measurement of self-reported emotional experience at the time of judgment (Lotto et al., 2014; Pletti et al., 2016; Sarlo et al., 2012), showing a low emotional involvement in incidental scenarios.

To date, several theoretical frameworks and empirical data on moral judgments for sacrificial and incidental dilemmas are available in the literature, in terms of moral acceptability, decision times, and emotional activation (Cushman et al., 2006; Greene et al., 2001, 2008; Lotto et al., 2014; Moore et al., 2008; Navarrete et al., 2012; Ugazio et al., 2012), typically applied to a wide variety of extreme and high-conflict events in heterogeneous settings. To date, only a very small number of these scenarios have been properly structured and tested to for the context of moral driving behavior, and mainly as an adaptation of the traditional trolley problem. Moreover, the majority of these cases are applied to autonomous transportation, whose social acceptance, perceived safety, and trust are still a work in progress (e.g., Ghazizadeh et al., 2012; Hengstler et al., 2016; Jing et al., 2020), leaping over the traditional non-autonomous driving level.

Indeed, although applying the traditional sacrificial dilemma structure to the driving activity—with or without a human driver—seems plausible, the significant amount of experimental research on the ethical perception of driver behavior has not yet provided a solid experimental validation.

The Present Study

Until today, relatively low attention has been paid to (a) the reliability of thought experiments as experimental tools in the context of traffic behavior and (b) how these scenarios should be structured. A variety of general rules has been settled concerning the structuring of text-based driving dilemmas, leading to a case-by-case adaptation of AV/driving dilemmas. In the present research, we decided to take a step back from the contemporary autonomous transportation application, with the aim of ensuring an experimental and structural validation of sacrificial manual driving dilemmas. This passage seems crucial to ensuring the possibility of applying the trolley problem in the traffic psychology framework, and consequently the reliability of the experimental conclusions in traffic and AV ethics obtained so far.

To this aim, we compared, in text-based form, a set of traditional sacrificial trolley problems (Lotto et al., 2014, as modified by Sarlo et al., 2012) with a novel set of trolley-like manual driving dilemmas. We hypothesized no differences between the two sets, in terms of decision time, moral judgment, emotional activation and moral acceptability. This result may be consistent with the idea of a negligible effect of the driving customization of the dilemmas on moral decision processes, confirming the validity of this experimental tool in on-road scenarios.

Furthermore, considering the direct involvement of the moral agent in the execution of the driving activity, the

potential role of the self-involvement risk factor in the decision process was also considered in the development of the new manual driving dilemmas set. In this regard, we expected a higher frequency of self-protective decisions in self-involvement dilemmas, which should be considered as more unpleasant and more arousing than dilemmas without individual risk. These results may be also considered as a replication of Lotto and colleagues (2014).

Through the proposed sacrificial driving scenarios, we provide a new specific set of incidental dilemmas, rigorously controlled for a number of psychological and linguistic confounding factors, on the basis of the DDE. In this context, the Italian normative values for the following variables were provided: (a) rates of participants' choices in each scenario; (b) decision times; (c) ratings of emotional valence and arousal experienced during the decision process, and (d) judgments of the moral acceptability of the two proposed driving behaviors.

Methods

Participants

A baseline equation assuming a medium effect size (Cohen's $d=0.25$) and a correlation of 0.50 among repeated measures, with a bidirectional hypothesis and an alpha error probability of 0.05 with 0.90 power, was tested before any data analysis with the G-power statistical software (Faul & Erdfelder, 1992). The system suggested 124 participants (31 per group). We recruited 152 participants (75 women, 1 unspecified) for the experiment. Each participant gave formal written consent prior to participation, which was voluntary and unremunerated. Mean age was 25.7 ($SD=5.48$, range = 18–57), and 69.08% were enrolled in university courses ($n=105$), with 40% matriculated in human sciences degree programs (e.g., psychology or sociology; $n=42$) and 21% in technical courses (e.g., engineering or mathematics; $n=22$). Most participants (95.4%) had held driver licenses ($n=145$) for 6.56 years on average ($SD=5.45$, range: 1–39). Almost all participants (99.3%, $n=151$) drove a maximum of 15,000 km per year, and only 4.60% were involved in a car accident in the prior 12 months ($n=7$). The study was approved by the local ethics committee (ID No.: 3514).

Stimuli

The set of stimuli was composed of 42 dilemmas: 40 sacrificial and 2 “filler” nonsacrificial moral scenarios (e.g., being dishonest). The two fillers, as well as 20 of the sacrificial dilemmas (“traditional dilemmas”), were selected from the validated set of Lotto et al., (2014, edited from Sarlo et al., 2012), whereas the remaining 20 scenarios were newly

developed as driving-type dilemmas. The new driving-type subset was developed as typical incidental dilemmas based on the DDE (Aquinas, 1952). We decided to focus our attention solely on incidental scenarios (i.e., killing one person as a foreseen but unintended consequence of saving many) because of the difficulty in imagining and producing plausible behind-the-wheel moral judgments in the classic instrumental version.

As with the traditional validated set, driving-type dilemmas are composed of a hypothetical moral text scenario and two possible resolutions: a deontological action and a utilitarian action. In both categories of dilemmas, the self-risk involvement factor was considered: in 10 scenarios per type, the agent was not involved as a potential victim of the accident (“other-involvement dilemmas”), whereas, in the remaining 10 scenarios, the utilitarian outcome resulted in protecting his own and other people’s lives, but sacrificing a single individual as an unintended and predicted side effect (“self-involvement dilemmas”) (Table 1). The driving-type self-involvement scenarios were structured consistently with the validated self-involvement set (Lotto et al., 2014), sharing the same self-sacrificial frame of more widely used moral dilemmas (i.e., see the “Crying Baby”, the “Plane Crash” and the “Sacrifice” scenarios, Greene et al., 2001, 2004, 2008).

The new driving-type set was controlled for several potential factors that may blur the ethical decision process (Awad et al., 2018; Lotto et al., 2014). These potential factors included (a) consideration of traffic rules and regulation to prevent a clear allocation of responsibilities; (b) avoidance of “leading” language and critical words (e.g., response bias; Loftus & Palmer, 1974); (c) avoidance of characterizing passengers and pedestrians in terms of gender, age, and personal relationship with the driver; (d) a constant 1:3 ratio between killed and saved characters; (e) limitation of the scenarios to a unique moral dilemma (i.e., killing or letting die); and (f) careful control of each scenario for the number of letters and words (see supplementary material: https://osf.io/cp285/?view_only=37c928fa5dc047d6905e40521a6a48dc).

Experimental Design and Procedure

The experimental task was programmed and distributed on Qualtrics survey software, and the complete experimental procedure had a medium duration of 26 min ($SD=7.26$ min).

All the participants were requested to complete the survey using a fixed device or a laptop, because of two reasons: the limited comparability of data when obtained through multiple device surveys (e.g. Krebs & Höhne, 2021) and to avoid that the participant performed the experiment while on board of a road vehicle.

For methodological reasons, participants were divided into four “lists” ($n=38$ per list), each of which was composed of 18 properly randomized dilemmas: eight traditional, eight driving-type, and the two filler scenarios. Four couples of these 18 dilemmas—one couple per category—were “list-specific,” thus only present in one particular list, whereas the remaining five couples or “anchors” (including fillers) were common between all the lists.

This experimental design was chosen for two reasons. First, because the intrinsic nature of the experimental activity requests a certain amount of effort and time for each scenario, the administration of all the dilemmas to all participants was considered too demanding and time consuming. A preliminary pilot study conducted on 12 participants confirmed that the 42-dilemma version was perceived as too long. In fact, the administration of 42 dilemmas required more than 1 h per participant. Given these results, the set of dilemmas to be administered was reduced to 18 per list. In this way, the experimental duration was halved, while scenarios were paired per category in each list. This approach granted the collection of a sufficient number of answers for each “list-specific” scenario (38), to obtain an adequate a priori statistical power.

The “anchor” dilemmas, selected from the normative scores of Sarlo et al. (2012) and pilot scores, had the aim of ensuring the reliability of each scenario in its own category. This was allowed by a preliminary analysis on non differences between anchors of the same category and between lists.

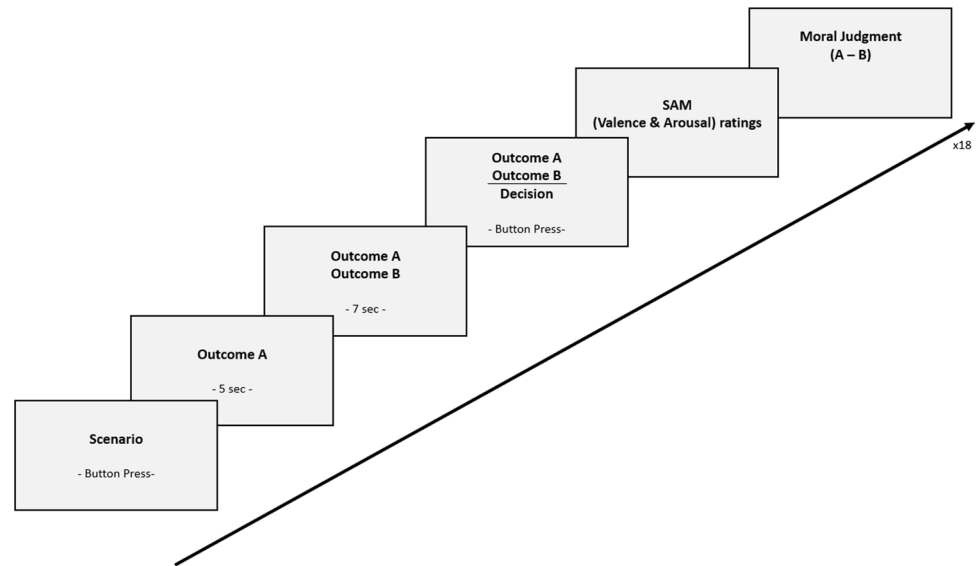
The experimental procedure is represented in Fig. 1. Due to the COVID-19 pandemic, participants completed the experiment online. Before starting, they were asked to complete the informed consent form and read instructions about the experimental session. Each dilemma was then presented as a text, in black type (font Arial, size 10) against a white background. For each dilemma, participants could read the scenario for as long as needed. Moving forward to the next slide, Outcome A (deontological action) was shown. After 5 s, Outcome B (utilitarian action) was added to the screen, and then, after 7 more seconds, options keys appeared allowing the participants to indicate the preferred outcome (for a detailed explanation of the procedure, see Palmiotti et al., 2020; Sarlo et al., 2012).

Subsequently, respondents rated their emotional state during the moral decision-making process using the Self-Assessment Manikin (SAM; Lang et al., 2008). The 9-point graphic scale was presented in two different slides, to evaluate valence (unpleasantness/pleasantness) and arousal (calm/activation), with higher scores indicating higher pleasantness and higher emotional arousal. Moving on to the last slide, participants were asked to rate on a 0–7 8-point scale ($7 = \text{completely acceptable}$) how morally acceptable the two proposed outcomes were. Decision times were recorded

Table 1 Sample non-customized, driving-type and filler dilemma (text translated from Italian)

Dilemma	Scenario	Resolutions
<i>Traditional Self-involvement (TS)</i>	A large man is leading you and two other explorers out of an underground cave on the west coast of Scotland. Due to its size, the man gets stuck in the narrow opening right at the exit of the cave. In a short time, the high tide will overwhelm you and everyone will drown except the man, who has his head out of the cave.	<p>A. Since the passage is blocked, you let the high tide overwhelm you and all three drowned</p> <p>B. You use the explosives you brought with you to widen the cave opening. You know that the man stuck will die, but you and the other two explorers will be able to escape.</p>
<i>Traditional Other-involvement (TO)</i>	You are a building worker who is maneuvering a crane on a building site. You have just started your day on the site, when you realize that the cable of the crane is about to break. Attached to the cable is an enormous steel beam which is directly above a crew of six who are working on the outside of a building in construction.	A. As there is no way to warn them, you let the cable break and the steel beam kills the six workers. B. You slightly move the crane arm to another area of the construction site. You know that two workers work there, who will be crushed by the steel beam and will die, but the six workers will be saved.
<i>Driving-type Self-involvement (DS)</i>	It is night, and you are driving your car. A violent storm has hit your city for a few hours; it is still raining, and the asphalt is slippery. You are approaching a traffic light when suddenly two cyclists cross the road right in front of you. On the right sidewalk, you see a pedestrian. You try to brake, but the tires slide on the asphalt.	A. You continue straight, running over the two cyclists, who will die. Your car will continue to slide on the asphalt, crashing against a nearby building, and you will die. B. You suddenly steer right. You know you will run over the pedestrian on the sidewalk, who will die, but your car will slow down in an open field and you, and the two cyclists will be unharmed.
<i>Driving-type Other-involvement (DO)</i>	You are driving your car on a panoramic road. A single cyclist is riding along the cycling path on your right, parallel to the roadway. As you drive on the road, you suddenly see three workers in the middle of the road removing a small obstacle. You are too close to them and do not have enough time to brake.	A. You continue straight, running over the three workers, who will die. B. You suddenly steer right. You know you will run over the single cyclist on the cycle path, who will die, but the three workers on the main road will be unharmed.
<i>Filler Non-sacrificial (FN)</i>	You were invited to a birthday by an acquaintance. You do not really want to go or spend a lot of money on the gift because of your superficial relationship with the person. You will find a branded sweater in excellent condition in a second-hand shop. When the package is opened, the birthday girl is embarrassed in front of such an important gift.	A. You smile satisfied and reassure the person by saying that an important occasion like a birthday deserved a very special gift. B. At the opening of the gift you immediately tell the person the truth about the sweater, saying that today's fashion trend is relaunching the use of branded second-hand clothes.

Fig. 1 Sequence of events in the experimental procedure for each dilemma



from the onset of the decision slide until the button press. The above experimental procedure was described step-by-step before the beginning of the experimental session, avoiding the presentation of an additional trial scenario.

Analysis

The statistical analysis was conducted in the R environment (version 4.0.4). In a preliminary step, the similarity of anchor dilemmas among lists was tested, looking for a nonsignificant interaction between the four lists and, respectively, dilemma type and specific anchor variables. We focused on outcome preferences (decision type, dichotomic), decision times, valence (1–9), and arousal (1–9). Given the nature of the collected data, linear mixed models and generalized linear models only for the dichotomous variable, with participants as random variables, were fitted to the data ('lme4' package in R; Bates et al., 2015).

Main data analysis focused on the complete experimental stimuli set, setting aside list grouping, and considering several potential predictors in six forward stepwise model comparisons ($M_1 - M_6$; further information on the procedure is available in the supplementary material). Decision type, decision time, valence, arousal, and moral acceptability for both the outcomes were set as dependent variables in different mixed-effect linear model comparisons, with participants as random variables. Post hoc pairwise comparisons were computed with the emmeans function from R package 'emmeans' (Lenth, 2020), setting Bonferroni correction as an adjustment method. Decision times were transformed in their natural logarithmic form (Lotto et al., 2014) and cleaned of outliers, applying a 98% acceptance interval. Filler dilemmas were excluded from the preliminary and

main data analysis and analyzed separately. Descriptive information can be retrieved in Table 3 and Table 4, and in greater detail in the supplementary materials, along with the dataset and the R script.

Results

Anchors

As predicted, no interaction was observed between lists and dilemma type ($\chi^2_9 = 4.40, p = 0.88$) and between lists and specific anchor ($\chi^2_{12} = 3.46, p = 0.99$) in terms of decision times. Similarly, no significant interactions in outcome preferences were found either between lists and dilemma type ($\chi^2_9 = 11.92, p = 0.21$) or between lists and specific anchor ($\chi^2_{12} = 10.38, p = 0.58$). Furthermore, no interaction was observed between lists and dilemma type in terms of valence ($\chi^2_9 = 3.58, p = 0.93$) and arousal ($\chi^2_9 = 5.33, p = 0.80$). Such lack of significance was also observed between lists and specific anchor, both in valence ($\chi^2_{12} = 2.54, p = 0.99$) and in arousal ($\chi^2_{12} = 14.04, p = 0.30$). These results are in line with the homogeneity hypothesis of the four lists, based on non-different responses between dilemma categories and specific anchors. They allowed consideration of the four lists as a whole in the subsequent statistical analysis of the complete dataset.

Complete Dataset

Table 2 shows the estimates from the computed models. Decision time model M_1 included risk involvement (self, other), dilemma type (traditional, driving-type),

Table.2 Beta estimates e p-values from M₁ to M₆

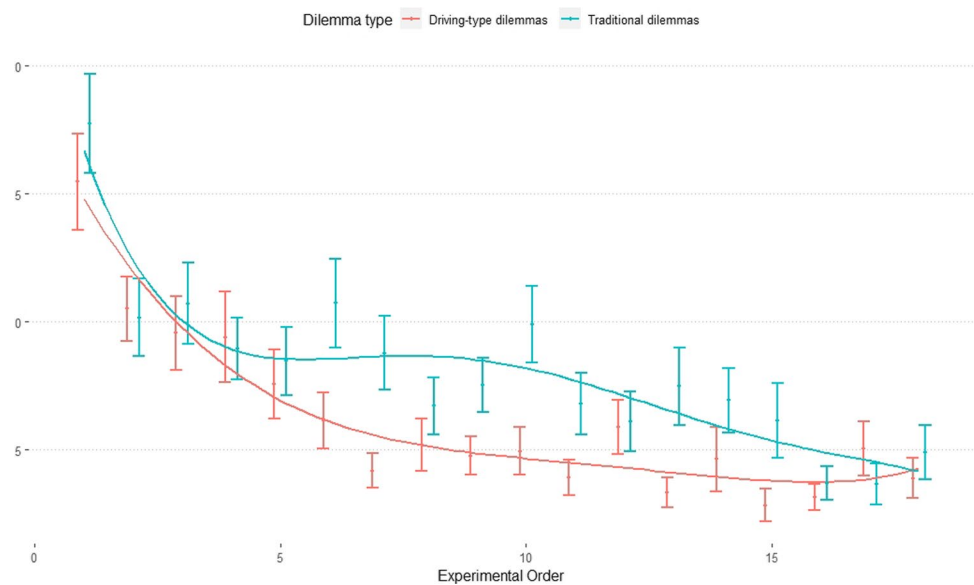
	N (%)	M ₁ Decision Time	M ₂ Decision Type	M ₃ Valence	M ₄ Arousal	M ₅ Morality Deont. Action	M ₆ Moral-ity Utilitarian Action
<i>Decision type</i>							
Deontological	15.00%	-	-	-	-	-	-
Utilitarian	85.00%	-0.37***	-	0.20**	-0.22*	-0.76*	0.74***
<i>Gender</i>							
Female	75	-	-	-	-	-	-
Male	75	-	-	0.61**	-0.98**	-	0.85***
NR	1	-	-	-	-	-	-
<i>Age</i>							
		-	-	-0.05*	-	-	-
<i>Degree Program (cluster)</i>							
Human Sciences	42	-	-	-	-	-	-
Cultural Studies	33	-	-	-0.77*	-	-	-
Technical Courses	22	-	-	-0.83*	-	-	-
NR	11	-	-	-0.40*	-	-	-
Not students	24	-	-	-0.49	-	-	-
Natural Sciences	2	-	-	-0.81	-	-	-
Humanistic Studies	18	-	-	-0.33	-	-	-
<i>Profession</i>							
Else	47	-	-	-	-	-	-
Students	105	-	-	-0.70**	-	-	-
<i>Risk Involvement</i>							
Self		-	-	-	-	-	-
Other	8	-0.35***	-	-	-	-	-
Experimental Order	8	-0.78***	0.43	0.08	-0.15*	-0.23***	0.11*
<i>Dilemma category</i>							
Traditional	8	-	-	-	-	-	-
Driving-Type	8	0.78***	-0.59***	0.07	0.04	0.40***	-0.07
<i>Dilemma Type</i>							
DS – DO		0.35***	-0.43	-0.08	0.15	0.23*	-0.11
DS – TS		-0.07	0.59***	-0.07	-0.04	-0.40*	0.07
DS – TO		-0.09	0.80**	-0.06	0.04	0.22*	-0.14
DO – TS		-0.42***	1.03***	0.00	-0.20*	-0.63*	0.19*
DO – TO		-0.45***	1.24***	0.01	-0.12	0.01	-0.02
TS – TO		0.03	0.21	0.01	0.08	0.62*	-0.21*
Emotional State		-	-	-	-0.28***	-	-
Valence		-	-	-0.22***	-	-	-
Arousal		-	-	-	-	-	20***
Moral Judgment		-	-	-	-	0.22***	-
<i>Deontological Outcome</i>							
<i>Utilitarian Outcome</i>							
R ² marg (R ² adj) *		0.15 (0.36)	0.05 (0.30)	0.18 (0.57)	0.11 (0.68)	0.10 (0.59)	0.12 (0.65)

Note: **p* < 0.05; ***p* < 0.01; ****p* < 0.001

Table.3 Mean and Standard Deviation of the dependent variables considered, divided by Dilemma Type (traditional, driving-type) and Risk Involvement (self, other involvement)

	Traditional	Driving-type	Other	Self
Decision Time (sec)	8.34 (11.14)	6.06 (8.88)	6.99 (10.36)	7.41 (9.66)
Decision Type: utilitarian	80.12%	89.07%	83.59%	85.60%
Valence	2.69 (1.72)	2.68 (1.64)	2.73 (1.64)	2.65 (1.72)
Arousal	5.84 (2.26)	5.80 (2.25)	5.79 (2.23)	5.86 (2.27)
Morality: Deontological action	2.01 (1.98)	1.74 (1.81)	1.68 (1.81)	2.07 (1.99)
Morality: Utilitarian action	2.29 (2.05)	2.26 (1.98)	2.32 (2.03)	2.23 (1.99)

Fig. 2 Smoothed curves plot with error bars plot representing means and standard errors for decision times (in seconds), divided by experimental order and dilemma type



experimental order (1 to 18), and decision type (deontological, utilitarian) as fixed effects, as well as the interaction between risk involvement and dilemma type. A significant increase in decision times was observed in dilemmas comprising self-involvement ($\chi^2_1 = 17.05$, $p < 0.001$), as in the case of deontological decisions ($\chi^2_1 = 38.62$, $p < 0.001$). As expected, slower decision times were observed at the beginning of the experiment, with an increase in velocity during the experimental time ($\chi^2_1 = 427.66$, $p < 0.001$; Fig. 2). At the same time, an overall effect of dilemma type—slower decisions in traditional dilemmas ($\chi^2_1 = 43.10$, $p < 0.001$)—and an interaction effect between dilemma type and risk involvement ($\chi^2_1 = 24.41$, $p < 0.001$) were observed. Post hoc pairwise comparisons showed quicker resolutions of driving-type scenarios not involving the participant as a potential victim than traditional did (TO–DO: $z = 8.13$; $p < 0.001$) and driving-type self-involvement scenarios did (DS–DO: $z = 6.41$, $p < 0.001$), indicating less cognitive conflict in this type of dilemma.

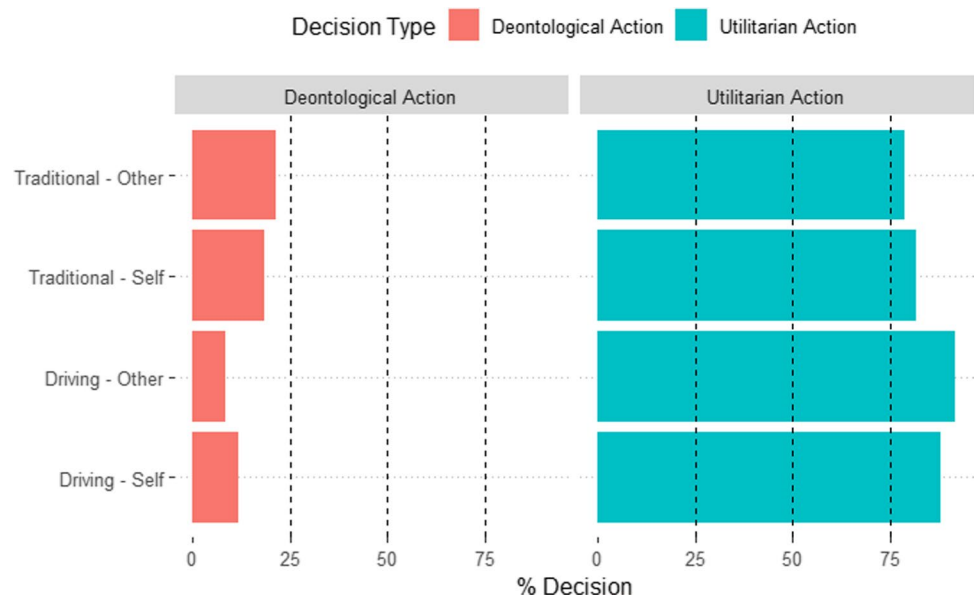
For the decision-type model M_2 , the binomial family distribution was set as a reference point for implementing a generalized mixed effects linear model, with subject as random variable. Following the comparison of the models, the one with risk involvement and dilemma type as fixed effects, and the interaction between them was selected. A higher percentage of utilitarian resolution approach was observed in driving scenarios (89%) when compared to traditional ones (80%, $\chi^2_1 = 47.34$, $p < 0.001$), whereas no risk involvement effect was detected. The interaction between involvement and dilemma type was significant ($\chi^2_1 = 6.29$, $p = 0.012$): driving-type dilemmas were solved more frequently through the utilitarian option, both when the decider's life was at risk (TS–DS: $z = -3.36$; $p = 0.004$) and when it was not (TO–DO: $z = -6.53$; $p < 0.001$; Fig. 3).

Focusing on valence, forward stepwise model comparisons led to an M_3 mixed effects linear model including risk involvement, dilemma type, decision type, arousal, gender, degree program (7 levels), profession (student, nonstudent), and age as fixed effects, as well as the interaction between risk involvement and dilemma type. No differences were observed in terms of risk involvement ($\chi^2_1 = 0.62$, $p = 0.430$), dilemma type ($\chi^2_1 = 0.50$, $p = 0.478$), and in the interaction between the two factors ($\chi^2_1 = 1.02$, $p = 0.312$), while lower scores of unpleasantness were detected when the deontological outcomes were selected ($\chi^2_1 = 8.08$, $p = 0.004$) and for higher activation scores ($\chi^2_1 = 174.44$, $p < 0.001$). Women indicated higher levels of unpleasantness than men did ($\chi^2_1 = 7.63$, $p = 0.022$), and older people ($\chi^2_1 = 5.68$, $p = 0.017$) indicated higher unpleasantness than the overall sample did. Interestingly, students also showed higher scores of unpleasantness during moral decision-making than non-students did ($\chi^2_1 = 5.57$, $p = 0.018$).

When arousal as dependent variable was selected, we computed a mixed linear model M_4 containing risk involvement, dilemma type, decision type, valence, and gender as fixed effects, as well as the interaction between risk involvement and dilemma type. As expected, higher activation matched dilemmas with personal involvement ($\chi^2_1 = 5.54$, $p = 0.018$), whereas no statistical significance was detected looking at dilemma type factor ($\chi^2_1 = 2.43$, $p = 0.118$) and its interaction with risk involvement ($\chi^2_1 = 0.45$, $p = 0.501$). Women showed higher arousal levels than men did ($\chi^2_1 = 13.48$, $p = 0.001$), and coherently with the observed negative relation between valence and arousal ($\chi^2_1 = 156.53$, $p < 0.001$), the selection of deontological options led to higher scores of the dependent variable.

Models M_5 and M_6 focused on the moral judgment concerning the two proposed outcomes, respectively

Fig. 3 Flipped bar chart of decision percentage frequencies, divided by involvement and dilemma type in rows and by decision type in columns (Deontological, Utilitarian)



deontological and utilitarian, independently from the chosen behavior. A dilemma effect on the deontological option (M_S) was observed, with lower scores judging driving-type dilemmas ($\chi^2_1 = 19.85, p < 0.001$), as well as a significant interaction effect between risk involvement and dilemma type ($\chi^2_1 = 17.87, p < 0.001$). In traditional self-involvement dilemmas, the deontological option received higher ratings of moral acceptability than the ones of the same typology without personal involvement (NCS–NCO: $z = 9.46, p < 0.001$) and then driving-type self-involvement scenarios (NCS–DTS: $z = 6.14, p < 0.001$). A coherent overall effect of Decision Type on the moral evaluation of both the moral outcomes was confirmed (deontological:

$\chi^2_1 = 107.64, p < 0.001$; utilitarian: $\chi^2_1 = 107.21, p < 0.001$; Fig. 4). Gender differences related only to the moral rating of utilitarian behavior were observed, with lower ratings of moral agreement expressed by women ($\chi^2_1 = 14.97, p < 0.001$).

Finally, four separate linear mixed effect models were performed for filler dilemmas on logarithmic decision times, selected outcome (generalized), valence, and arousal, with subject as random variable and gender (male, female), profession (student, otherwise), degree program (7 levels), and decision type (immoral, moral behavior). No relevant effects were observed in any of the four models. Descriptive information can be retrieved in Table 3 and Table 4.

Fig. 4 Error bars plot representing means and standard errors of participants' moral evaluations, divided by involvement and dilemma type in rows (DS: Driving-type Self; DO: Driving-type Other; TS: Traditional Self; TO: Traditional Other), and decision type in columns (deontological action, left; utilitarian action, right)



Table 4 Mean and Standard Deviation of the dependent variables considered, divided by the interaction between Dilemma Type and Risk Involvement

	Traditional Other	Traditional Self	Driving-type Other	Driving-type Self
Decision Time (sec)	8.80 (12.01)	7.87 (10.21)	5.17 (8.64)	6.96 (9.11)
Decision Type: utilitarian	77.02%	81.51%	91.66%	87.98%
Valence	2.70 (1.63)	2.68 (1.80)	2.75 (1.65)	2.61 (1.64)
Arousal	5.80 (2.22)	5.88 (2.29)	5.77 (2.25)	5.83 (2.25)
Morality: Deontological action	1.76 (1.87)	2.25 (2.10)	1.60 (1.75)	1.88 (1.87)
Morality: Utilitarian action	2.32 (2.10)	2.27 (1.99)	2.32 (1.97)	2.20 (1.98)

Discussion

The present study aimed to investigate the suitability of the trolley-problem research tool when applied to moral driving behavior in on-road scenarios. To this end, a selection of validated incidental dilemmas was compared with newly developed driving-type dilemmas, structured based on the DDE. The behavioral choice, decision times, emotional state, and the moral acceptability of the two proposed behaviors were analyzed, and the corresponding normative scores were reported.

Results showed that both sets of dilemmas led to more frequent utilitarian than deontological outcome selections, as well as to comparable scores of unpleasantness and arousal and low moral acceptability regardless of the assumed decision, suggesting that similar decision processes were involved during the dilemma resolution, which also seems resistant to demographic, educational, and gender differences (Navarrete et al., 2012; Palmiotti et al., 2020). Therefore, we can lend credence to a fundamental role of structure over the context in the interpretation of the trolley problem (Schein, 2020). A “structure-based” interpretation of moral dilemmas should be supported by no differences in moral judgments and emotional activation between traditional and driving incidental scenarios. On the contrary, score differences would represent different interpretations of the moral problems, underlying the importance of each scenario’s contextualization in its consequent perception (“context-based” hypothesis). In the present research, the behavioral attitudes shown towards the proposed scenarios were not dependent on their customization, but the driving-type seemed to enhance the utilitarian moral code strongly. In fact, regardless of the potential personal risk, resolutions were faster for on-road trolley problems than for traditional dilemmas. This phenomenon may suggest that on-road storylines are easier to contextualize in daily challenges than traditional moral dilemmas, which are typically applied to extreme and often unrealistic circumstances. Our belief is that when moral judgments and behaviors become more lifelike, cognitive demands decrease (Conway & Gawronski, 2013; Schein, 2020; Sütfield et al., 2019), allowing the individual’s moral inclination to emerge more automatically

and effectively. Furthermore, this result lends new credence to the criticism of low-plausibility dilemmas as a serious confounder of moral judgment (e.g., Bauman et al., 2014; Gold et al., 2014). For example, when moral dilemmas are detected as scarcely plausible, Körner and colleagues (2019) observed a systematic increase of deontological judgments, triggering a potential distortion in the examination of moral cognition. The on-road driving activity seems a simple example of this phenomenon, as the utilitarian approach to driving dilemmas emerged faster and sharper.

In the context of the dual-process theory, previous studies focused on the role of time pressure on moral judgment, discovering a negative effect on the proportion of the utilitarian resolution to sacrificial dilemmas (Jaquet & Cova, 2021; Suter & Hertwig, 2011; Trémolière & Bonnefon, 2014; Trémolière et al., 2017). This evidence seems helpful in explaining the high percentage of utilitarian behaviors in the presented study, where no time pressure was planned for each dilemma. Nonetheless, Tinghög et al. (2016) did not observe an effect of time manipulation on moral judgment, contrasting our interpretation. The topic deserves a more in-depth analysis in future studies.

Interestingly, results also show an overall improved responsiveness to sacrificial dilemmas during the course of the experiment. The decreasing trend in reaction times during the completion of iterated trolley problem experiments is a slightly addressed issue in the field of moral psychology. In the present research, the frequent selection of the utilitarian moral code and the constant reduction of decision times during the whole task opens a new discussion on the potential risks of the iterated approach. Traditionally, one important methodological factor that determines the statistical power of experimental results is the definition of a sufficient number of trials (e.g., Baker et al., 2020; Forrester, 2015; Lerche et al., 2017). Solving a moral dilemma is allegedly a conflictual task to face (e.g., Broeders et al., 2011) in terms of time and cognitive effort, and the attempt to operationalize its experimental load on the participant does not seem trivial. In this context, presenting more moral scenarios than needed may have a detrimental effect, potentially increasing the distance from the true value and so from a proper definition of the individual moral code. This noise effect may be due

to several causes that can reduce participants' commitment to the task, all conceivably related to the standard within-subject experimental design (e.g., high number of trials, cognitive fatigue, boredom, emotional activation, desire to finish the experiment). Future iterated moral dilemma applications are needed to address this methodological question.

A structure-based hypothesis concerning the trolley problem is also supported by the results concerning the emotional state experienced during decision-making, indicating comparable unpleasantness and activation between the two dilemma types. The inverse relation between valence and arousal ratings (i.e., more unpleasantness associated with higher activation) for both traditional and driving-type dilemmas is consistent with previous findings on incidental scenarios (Lotto et al., 2014). As in traditional sacrificial dilemmas, emotional activation also seems to play a pivotal role in the on-road moral decision process when human lives are at stake (e.g., Cushman & Greene, 2012; Cushman et al., 2012; Haidt, 2007; Szekely & Miu, 2015). In the present study, the emotional state during driving scenarios is qualitatively and quantitatively comparable to the one observed during traditional dilemmas.

Predictably, data also showed how both the available outcomes were constantly judged as scarcely moral. Nevertheless, it is also possible to observe how the chosen action was consistently judged as less immoral than the alternative option, independently from the selected behavior and the dilemma type. This perception can be interpreted as a confirmation bias of the decision-maker, defined as the tendency to interpret evidence selectively to reinforce one's current beliefs or decisions (Nickerson, 1998). This result confirms the widely acknowledged dissociation between moral choice and moral judgment (Francis et al., 2016; Tassy et al., 2013), also observed in the context of autonomous transportation. Indeed, Bonnefon et al. (2016) highlighted the so-called "social dilemma of AVs" as the individual moral disagreement on utilitarian AVs, which are considered morally rightful for the community but unfair and unappealing for the individuals as passengers. As in our results, a mismatch between preferred behavior and its moral evaluation is apparent.

Surprisingly, the self-risk involvement factor did not have a significant effect on moral judgment and emotional activation. Manifold studies have observed the fundamental role of life-threatening scenarios in the development of a proper moral decision (Huebner & Hauser, 2011; Lotto et al., 2014; Moore et al., 2008; Petrinovich et al., 1993; Sütfield et al., 2019), recognizing the direct proportion between utilitarian decisions and the potential number of lives saved (Bergmann et al., 2018; Bonnefon et al., 2016). This argument has not only theoretical but also practical implications: a threat to life is certainly a conflictual factor to cope with for a driver, both when driving a vehicle and when simply carried by an

autonomous one. This leads us to recognize the importance of stressing further this topic in future applications.

This study has some potential limitations. The adaptation of the typical incidental dilemma to the context of manual driving results in a new perception of the individual agency in the available moral options. Traditionally, a proactive utilitarian decision is compared to a passive deontological one, in which the agent's inaction leads to a costly coincidental causal mechanism. When the agent is placed with her/his hands on the wheel, his endorsement of the passive action can be legitimately perceived as a deliberate decision to run onto several pedestrians, with or without self-sacrifice. Although the deontological option of the driving-type dilemmas shares the traditional dichotomy between intervening and taking no action against an imminent threat, this factor has to be considered.

Additionally, in order to be consistent with the dilemma structure of Lotto and colleagues (2014) and with other previous procedures (Palmiotti et al., 2020; Sarlo et al., 2012), the presentation of the two moral outcomes was not counter-balanced between and within participants. This decision was based on the need to postpone, as requested, the reasoning process at the presentation of the decision slide. The utilitarian option depicts a predictable loss in order to achieve a collective benefit (e.g., save more people). Using this design, we wanted to avoid the risk of anticipating the decision on the only basis of the proactive utilitarian alternative, so on the individual's disposition to harm someone for a greater good.

Further applications may be helpful in deepening these results in future researches: aiming for a wider cross-cultural validation (e.g., Awad et al., 2018; Di Stasi et al., 2020), strengthening the comparison between experimental modalities (e.g., Virtual Reality; see Sütfield et al., 2017, 2019; Vankov & Jankovszky, 2021), controlling moral judgments for individual and situational factors (e.g. baseline affective state, online emotional state or personality traits, see Klenk, 2021), as well as deepening the role of social distance in determining moral judgment (Hofer et al., 2020), the causality of the agent in the dramatic consequences of the proposed outcomes (Phillips & Shaw, 2015), or the tendency to self-protection in pursuing the utilitarian behavior (i.e., framing the self-protective option both in the utilitarian and in the non-utilitarian outcome).

In conclusion, through the present research, we brought new evidence on the suitability of the incidental dilemma tool in the investigation of moral driving behaviors, suggesting the possibility that on-road contextualization leads to faster resolution of dilemmas with comparable emotional activation. Regardless of their contextualization, these kinds of hypothetical scenarios are frequently employed to unveil ethical issues relevant to everyday life. Although these dilemmas may seem sometimes radical and unrealistic,

it helps us untie the cognitive and emotional mechanisms behind moral decisions. Nonetheless, moral judgment behind the wheel appears to be closer to our daily challenges, as opposed to the traditionally broader moral scenarios. We believe that researchers in traffic behavior and moral decision-making may benefit from this new driving-type dilemma set, to easily approach the complex task of developing and testing realistic moral scenarios in the context of autonomous and nonautonomous transportation.

Acknowledgements This work was carried out within the scope of the project "use-inspired basic research", for which the Department of General Psychology of the University of Padova has been recognized as "Dipartimento di Eccellenza" by the Ministry of University and Research.

Authors' Contributions Conceptualization: Giovanni Bruno; Methodology: Giovanni Bruno, Michela Sarlo, Lorella Lotto, Nicola Cellini, Simone Cutini, Andrea Spoto; Formal analysis and investigation: Giovanni Bruno, Andrea Spoto; Writing—original draft preparation: Giovanni Bruno; Writing—review and editing: Michela Sarlo, Lorella Lotto, Nicola Cellini, Simone Cutini, Andrea Spoto; Resources: Giovanni Bruno, Michela Sarlo, Lorella Lotto; Supervision: Michela Sarlo, Lorella Lotto, Simone Cutini, Andrea Spoto.

Data Availability The data that support the findings of this study are openly available in Open Science Framework at: https://osf.io/cp285/?view_only=37c928fa5dc047d6905e40521a6a48dc

Declarations

Conflicts of Interest The authors declare no conflict of interest.

Ethics Approval The study was approved by the local ethics committee (ID No.: 3514).

Consent to Participate Informed consent was obtained from all individual participants included in the study.

References

- Aquinas, T. (1952). The summa theologica (fathers of the english dominican province, trans.). In W. Benton (Series Ed.), *Great Books Series: Vol. 19. Chicago: Encyclopedia Britannica, Inc.* (Original work published 1274).
- Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J. F., & Rahwan, I. (2018). The Moral Machine experiment. *Nature*, *563*(7729), 59–64. <https://doi.org/10.1038/s41586-018-0637-6>
- Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2020). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *Psychological Methods*.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Baum, S. D. (2020). Social choice ethics in artificial intelligence. *AI & SOCIETY*, 1–12.
- Bauman, C. W., McGraw, A. P., Bartels, D. M., & Warren, C. (2014). Revisiting external validity: Concerns about trolley problems and other sacrificial dilemmas in moral psychology. *Social and Personality Psychology Compass*, *8*(9), 536–554.
- Bentham, J. (1781). An introduction to the principles of morals and legislation. McMaster University Archive for the History of Economic Thought.
- Bergmann, L. T., Schlicht, L., Meixner, C., König, P., Pipa, G., Boshammer, S., & Stephan, A. (2018). Autonomous vehicles require socio-political acceptance—An empirical and philosophical perspective on the problem of moral decision making. *Frontiers in behavioral neuroscience*, *12*(31).
- Bonnefon, J. F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles. *Science*, *352*(6293), 1573–1576.
- Borg, J. S., Hynes, C., Van Horn, J., Grafton, S., & Sinnott-Armstrong, W. (2006). Consequences, action, and intention as factors in moral judgments: An fMRI investigation. *Journal of Cognitive Neuroscience*, *18*, 803–817.
- Broeders, R., Van den Bos, K., Müller, P. A., & Ham, J. (2011). Should I save or should I not kill? How people solve moral dilemmas depends on which rule is most accessible. *Journal of Experimental Social Psychology*, *47*(5), 923–934.
- Chan, M., & Singhal, A. (2015). Emotion matters: Implications for distracted driving. *Safety Science*, *72*, 302–309.
- Conway, P., & Gawronski, B. (2013). Deontological and utilitarian inclinations in moral decision making: A process dissociation approach. *Journal of Personality and Social Psychology*, *104*(2), 216–235. <https://doi.org/10.1037/a0031021>
- Cushman, F., & Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social Neuroscience*, *7*, 269–279. <https://doi.org/10.1080/17470919.2011.614000>
- Cushman, F., Young, L., & Hauser, M. (2006). The role of conscious reasoning and intuition in moral judgment: Testing three principles of harm. *Psychological Science*, *17*(12), 1082–1089. <https://doi.org/10.1111/j.1467-9280.2006.01834.x>
- Cushman, F., Gray, K., Gaffey, A., & Mendes, W. B. (2012). Simulating murder: The aversion to harmful action. *Emotion*, *12*(1), 2–7. <https://doi.org/10.1037/a0025071>
- De Sio, F. S. (2017). Killing by autonomous vehicles and the legal doctrine of necessity. *Ethical Theory and Moral Practice*, *20*(2), 411–429.
- Di Stasi, L. L., Diaz-Piedra, C., Morales, J. M., Kurapov, A., Tagliabue, M., Bjärtå, A., ... & Catena, A. (2020). A cross-cultural comparison of visual search strategies and response times in road hazard perception testing. *Accident Analysis & Prevention*, *148*, 105785.
- Faul, F., & Erdfelder, E. (1992). *GPOWER: A priori, post-hoc, and compromise power analyses for MS-DOS [Computer program]*. Bonn University, Department of Psychology.
- Foot, P. (1978). *The problem of abortion and the doctrine of double effect*. In *Virtues and vices*. Blackwell.
- Forrester, S. E. (2015). Selecting the number of trials in experimental biomechanics studies. *International Biomechanics*, *2*(1), 62–72.
- Francis, K. B., Howard, C., Howard, I. S., Gummerum, M., Ganis, G., Anderson, G., & Terbeck, S. (2016). Virtual morality: Transitioning from moral judgment to moral action? *PLoS ONE*, *11*(10). <https://doi.org/10.1371/journal.pone.0164374>
- Frison, A. K., Wintersberger, P., & Rieger, A. (2016). First person trolley problem: Evaluation of drivers' ethical decisions in a driving simulator. *AutomotiveUI 2016 - 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Adjunct Proceedings*, 117–122. <https://doi.org/10.1145/3004323.3004336>
- Ghazizadeh, M., Lee, J. D., & Boyle, L. N. (2012). Extending the Technology Acceptance Model to assess automation. *Cognition, Technology & Work*, *14*(1), 39–49.
- Gkartzonikas, C., & Gkritza, K. (2019). What have we learned? A review of stated preference and choice studies on autonomous

- vehicles. *Transportation Research Part c: Emerging Technologies*, 98, 323–337.
- Gold, N., Pulford, B. D., & Colman, A. M. (2014). The outlandish, the realistic, and the real: Contextual manipulation and agent role effects in trolley problems. *Frontiers in Psychology*, 5, 35.
- Graham, J., Meindl, P., Beall, E., Johnson, K. M., & Zhang, L. (2016). Cultural differences in moral judgment and behavior, across and within societies. *Current Opinion in Psychology*, 8, 125–130.
- Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, 293(5537), 2105–2108. <https://doi.org/10.1126/science.1062872>
- Greene, J. D., Nystrom, L. E., Engell, A. D., Darley, J. M., & Cohen, J. D. (2004). The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44(2), 389–400. <https://doi.org/10.1016/j.neuron.2004.09.027>
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, 107(3), 1144–1154. <https://doi.org/10.1016/j.cognition.2007.11.004>
- Haboucha, C. J., Ishaq, R., & Shiftan, Y. (2017). User preferences regarding autonomous vehicles. *Transportation Research Part c: Emerging Technologies*, 78, 37–49.
- Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316, 998–1002. <https://doi.org/10.1126/science.1137651>
- Hengstler, M., Enkel, E., & Duelli, S. (2016). Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change*, 105, 105–120.
- Hidalgo, C. A., Orghian, D., Canals, J. A., De Almeida, F., & Martín, N. (2021). *How Humans Judge Machines*. MIT Press.
- Hofer, M., Tamborini, R., & Ryffel, F. A. (2020). Between a rock and a hard place: The role of moral intuitions and social distance in determining moral judgments of an agent in a moral dilemma. *Journal of Media Psychology: Theories, Methods, and Applications*.
- Huebner, B., & Hauser, M. D. (2011). Moral judgments about altruistic self-sacrifice: When philosophical and folk intuitions clash. *Philosophical Psychology*, 24, 73–94. <https://doi.org/10.1080/09515089.2010.534447>
- Jaquet, F., & Cova, F. (2021). Beyond moral dilemmas: The role of reasoning in five categories of utilitarian judgment. *Cognition*, 209, 104572.
- Jing, P., Xu, G., Chen, Y., Shi, Y., & Zhan, F. (2020). The determinants behind the acceptance of autonomous vehicles: A systematic review. *Sustainability*, 12(5), 1719.
- Kallioinen, N., Pershina, M., Zeiser, J., NosratNezami, F., Pipa, G., Stephan, A., & König, P. (2019). Moral Judgements on the Actions of Self-Driving Cars and Human Drivers in Dilemma Situations From Different Perspectives. *Frontiers in Psychology*, 10(November), 1–15. <https://doi.org/10.3389/fpsyg.2019.02415>
- Kant, I. (1785). *Groundwork of the metaphysics of morals*.
- Klenk, M. (2021). The influence of situational factors in sacrificial dilemmas on utilitarian moral judgments. *Review of Philosophy and Psychology*, 1–33.
- Körner, A., Joffe, S., & Deutsch, R. (2019). When skeptical, stick with the norm: Low dilemma plausibility increases deontological moral judgments. *Journal of Experimental Social Psychology*, 84, 103834.
- Krebs, D., & Höhne, J. K. (2021). Exploring scale direction effects and response behavior across PC and smartphone surveys. *Journal of Survey Statistics and Methodology*, 9(3), 477–495.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). International affective picture system (IAPS): affective ratings of pictures and instruction manual. *Technical Report A-8*, University of Florida, Gainesville, FL.
- Lenth, R. (2020). emmeans: Estimated marginal means, aka least-squares means (Version 1.5.2–1) [R package].
- Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods*, 49(2), 513–537.
- Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of verbal learning and verbal behavior*, 13(5), 585–589.
- Lotto, L., Sarlo, M., & Manfrinati, A. (2014). A New Set of Moral Dilemmas: Norms for Moral Acceptability, Decision Times, and Emotional Salience. *Journal of Behavioral Decision Making*, 16(20), 6513–6525. <https://doi.org/10.1002/bdm>
- Martí-Vilar, M., Escrig-Espuig, J. M., & Merino-Soto, C. (2021). A systematic review of moral reasoning measures. *Current Psychology*, 1–15.
- Meder, B., Fleischhut, N., Krumnau, N., & Waldmann, M. R. (2019). How should autonomous cars drive? A preference for defaults in moral judgments under risk and uncertainty. *Risk Analysis*, 39(2), 295–314.
- Moore, A. B., Clark, B. A., & Kane, M. J. (2008). Who shalt not kill? Individual differences in working memory capacity, executive control, and moral judgment. *Psychological Science*, 19(6), 549–557.
- Nasello, J. A., Dardenne, B., Blavier, A., & Triffaux, J. M. (2021). Does empathy predict decision-making in everyday trolley-like problems? *Current Psychology*, 1–14.
- Navarrete, C. D., McDonald, M. M., Mott, M. L., & Asher, B. (2012). Virtual morality: Emotion and action in a simulated three-dimensional “trolley problem.” *Emotion*, 12(2), 364–370. <https://doi.org/10.1037/a0025561>
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220.
- Özkan, T., Lajunen, T., Chliaoutakis, J. E., Parker, D., & Summala, H. (2006). Cross-cultural differences in driving behaviours: A comparison of six countries. *Transportation Research Part f: Traffic Psychology and Behaviour*, 9(3), 227–242.
- Palmiotti, G. P., Del PopoloCristaldi, F., Cellini, N., Lotto, L., & Sarlo, M. (2020). Framing the outcome of moral dilemmas: Effects of emotional information. *Ethics and Behavior*, 30(3), 213–229. <https://doi.org/10.1080/10508422.2019.1607348>
- Petrinovich, L., O’Neill, P., & Jorgensen, M. (1993). An empirical study of moral intuitions: Toward an evolutionary ethics. *Journal of Personality and Social Psychology*, 64(3), 467.
- Phillips, J., & Shaw, A. (2015). Manipulating morality: Third-party intentions alter moral judgments by changing causal reasoning. *Cognitive Science*, 39(6), 1320–1347.
- Pigeon, C., Alauzet, A., & Paire-Ficout, L. (2021). Factors of acceptability, acceptance and usage for non-rail autonomous public transport vehicles: A systematic literature review. *Transportation Research Part f: Traffic Psychology and Behaviour*, 81, 251–270.
- Pletti, C., Lotto, L., Tasso, A., & Sarlo, M. (2016). Will I regret it? Anticipated negative emotions modulate choices in moral dilemmas. *Frontiers in Psychology*, 7(DEC), 1–15. <https://doi.org/10.3389/fpsyg.2016.01918>
- Samuel, S., Yahoodik, S., Yamani, Y., Valluru, K., & Fisher, D. L. (2020). Ethical decision making behind the wheel – A driving simulator study. *Transportation Research Interdisciplinary Perspectives*, 5, 100147. <https://doi.org/10.1016/j.trip.2020.100147>
- Sarlo, M., Lotto, L., Manfrinati, A., Rumiati, R., Gallicchio, G., & Palomba, D. (2012). Temporal dynamics of cognitive–Emotional interplay in moral decision-making. *Journal of Cognitive Neuroscience*, 24(4), 1018–1029. https://doi.org/10.1162/jocn_a_00146
- Schein, C. (2020). The Importance of Context in Moral Judgments. *Perspectives on Psychological Science*, 15(2), 207–215. <https://doi.org/10.1177/1745691620904083>

- Shariff, A., Bonnefon, J. F., & Rahwan, I. (2017). Psychological roadblocks to the adoption of self-driving vehicles. *Nature Human Behaviour*, *1*(10), 694–696. <https://doi.org/10.1038/s41562-017-0202-6>
- Suter, R. S., & Hertwig, R. (2011). Time and moral judgment. *Cognition*, *119*(3), 454–458. <https://doi.org/10.1016/j.cognition.2011.01.018>
- Sütfeld, L. R., Gast, R., König, P., & Pipa, G. (2017). Using virtual reality to assess ethical decisions in road traffic scenarios: Applicability of value-of-life-based models and influences of time pressure. *Frontiers in Behavioral Neuroscience*, *11*, 122.
- Sütfeld, L. R., Ehinger, B. V., König, P., & Pipa, G. (2019). How does the method change what we measure? Comparing virtual reality and text-based surveys for the assessment of moral decisions in traffic dilemmas. *PLoS ONE*, *14*(10), 1–14. <https://doi.org/10.1371/journal.pone.0223108>
- Szekely, R. D., & Miu, A. C. (2015). Incidental emotions in moral dilemmas: The influence of emotion regulation. *Cognition and Emotion*, *29*(1), 64–75.
- Tassy, S., Oullier, O., Mancini, J., & Wicker, B. (2013). Discrepancies between judgment and choice of action in moral dilemmas. *Frontiers in Psychology*, *4*(MAY), 1–8. <https://doi.org/10.3389/fpsyg.2013.00250>
- Thomson, J. (1985). The trolley problem. *Yale Law Journal*, *94*(6), 1395–1415.
- Thomson, J. J. (2008). Turning the trolley. *Philosophy & Public Affairs*, *36*(4), 359–374.
- Tinghög, G., Andersson, D., Bonn, C., Johannesson, M., Kirchler, M., Koppel, L., & Västfjäll, D. (2016). Intuition and moral decision-making—the effect of time pressure and cognitive load on moral judgment and altruistic behavior. *PLoS ONE*, *11*(10), 1–19. <https://doi.org/10.1371/journal.pone.0164012>
- Trémolière, B., & Bonnefon, J. F. (2014). Efficient kill–save ratios ease up the cognitive demands on counterintuitive moral utilitarianism. *Personality and Social Psychology Bulletin*, *40*(7), 923–930.
- Trémolière, B., De Neys, W., & Bonnefon, J. F. (2017). Reasoning and moral judgment: A common experimental toolbox. In L. J. Ball & V. A. Thompson (Eds.), *International handbook of thinking and reasoning* (pp. 575–589). Psychology Press.
- Ugazio, G., Lamm, C., & Singer, T. (2012). The role of emotions for moral judgments depends of the type of emotion and moral scenario. *Emotion*, *12*, 579–590. <https://doi.org/10.1037/a0024611>
- Vankov, D., & Jankovszky, D. (2021). Effects of using headset-delivered virtual reality in road safety research: A systematic review of empirical studies. *Virtual Reality & Intelligent Hardware*.
- Goodall, N. J. (2014). Ethical decision making during automated vehicle crashes. *Transportation Research Record*, *2424*(1), 58–65.

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