Hurting others vs hurting myself, a dilemma for our autonomous vehicle

Miguel Luzuriaga†, Antonio Heras‡, and Oliver Kunze†

October, 2019

Abstract.

Autonomous Vehicles (AVs) will soon be a reality on our roads. Up to now, their standard behaviour is predetermined since it is the result of programmed algorithms. Under risky traffic situations, however, they will face dilemmas such as, deciding between affecting the passenger or affecting others. Thus, in this experiment, we investigate how people solve a reframed version of the well-known Trolley-problem (Foot, 1967) under two conditions: when subjects actually drive a vehicle simulator, versus when they solve the same dilemma by programming a hypothetical AV. In both settings, the participants’ decisions have real monetary consequences, which affect others and themselves. Our Probit models indicate that subjects who program an AV and who are more cautious in terms of speed are less (more) likely to sacrifice a pedestrian (themselves) compared to those who actually drive and prefer a higher driving speed. Moreover, we find that the subjects’ choices are associated with risk aversion but not with moral beliefs or loss aversion. Implications of the driving vs programming discrepancies for the design of AVs algorithms are discussed.

Keywords: Experiments, Ethical dilemmas, Trolley Problem, Risk-aversion, Social Preferences

JEL Codes: C91, D81, D91, R41

* We would like to thank Florian Schlatterer, Bernhard Schick and participants at the CUM Workshop at the Complutense University of Madrid, 2017. Financial support from the Neu-Ulm University and TBS (Technology Network Bavaria-Swabia) is greatly appreciated. We also thank Wolf-Peter Maier, Johanna Hehr, and Alena Kotelnikov for their assistance in implementing the experiment.
† Research Center Allgäu, Kempten University of Applied Sciences, Germany, miguel.luzuriaga@hs-kempten.de
‡ Complutense University of Madrid, Spain, aheras@ccee.ucm.es
§ Neu-Ulm University of Applied Sciences, Germany, oliver.kunze@hs-neu-ulm.de
1. Introduction

Autonomous Vehicles (AVs) will soon be a reality on our roads. Leading technological companies such as Google, together with the most important car builders, are investing a huge amount of talent and effort in order to design safe and reliable vehicles able to transport goods and persons with no need of human drivers. No doubt, the new AVs will be far safer than the classical human-driven cars and trucks: the automatic drivers will never be victims of somnolence or distraction, will not consume alcohol or drugs, and will monitor all the important aspects of their environment, such as the weather conditions and the presence or absence of pedestrians and other vehicles. Since AVs will always be alert, and provided that possible software errors and failures will continuously be reduced to a minimum, the substitution of the human drivers by these new automatic devices will lead to a drastic reduction of traffic accidents and human casualties. Admittedly, AVs will help us to save money, time and pain.

The scientists and engineers designing and projecting these devices still have to overcome important technological challenges. Nevertheless, one of the important problems they face is not only a technical issue, and has important legal and philosophical connections. This problem is related to the behaviour of the AVs when faced to inevitable accidents. Because, no matter how big the risk reduction will be, the AVs are human creatures and therefore will still be involved in traffic accidents. The problem is that the behaviour of the AV is always predetermined, since it is the result of following some programming rules. It is not the human driver (who in fact does not exist) but the engineers who have to decide how the AVs behave in every possible traffic situation. The engineers have to decide in advance whether the most important goal must be to protect the passengers by all means, or minimize the total number of casualties, or any alternative objective. This is not an engineering or technical issue, but an ethical problem, that raises important moral dilemmas related to the distribution of the pains and evils in society. Even the governments are beginning to worry about this issue. In Germany, a national ethics committee for automated and connected driving appointed by the government has established that “the decision that has to be taken is whether the licensing of automated driving systems is ethically justifiable or possibly even imperative. If these systems are licensed – and it is already apparent that this is happening at international level – everything hinges on the conditions in which they are used and the way in which they are designed” (Luetge 2017).
Let us assume an inevitable accident: how should the AV behave? Probably it should protect the lives of the passengers to a certain extent. Many people would agree on this. But, what about the distribution of casualties outside the vehicle? Should the AV prefer to sacrifice the life of (say) an innocent pedestrian in order to save the lives of several (more than one) passengers of other vehicles on the road? In other words: should we have ‘utilitarian’ AVs programmed to minimize the total number of casualties when involved in traffic accidents? To some, but not all, this seems to be a sensible solution. In fact, it is difficult to reach a consensus on this question. Many years ago, when nobody could even dream of the existence of the AVs, moral philosophers designed a mental or thought experiment, the so-called ‘Trolley Problem’, which encapsulated the main features of this moral dilemma. In the basic version of this dilemma, an uncontrolled trolley heads towards a number of people (more than one) that are working on the track and will kill all of them, unless somebody (us?) pulls a lever and diverts the trolley onto a side track. In this last case, only one person (a different person) will be killed. Should we pull the lever and kill one in order to save several? This dilemma and its variations were the subject of lively discussions among the philosophers, who could not come to a solution acceptable for everybody. The problem was introduced by Philippa Foot in 1967, and since then has been a subject of controversy (see, among many others, Foot (1967), Thomson (1976, 1985), Kamm (1989), Unger (1996), and Edmonds (2013), Bruers & Braeckman (2014)).

This classical ethical version of the Trolley Problem has mutated and the old trolleys have revived under the new guise of AVs. In this new version, the person on the side track is a pedestrian standing on the sidewalk. But the dilemma remains the same: should the AV crash against several people on the road, killing them, or should it swerve and kill only one person (the pedestrian on the sidewalk)? Should the AV behave in a different way if the lives of the passengers are threatened instead of the lives of pedestrians outside the vehicle? In the recent years, a number of surveys has been conducted to uncover the preferences of real people in this basic dilemma and also in more complicated related problems. Not surprisingly, the results are not conclusive.

A large proportion of the empirical evidence on AVs has been obtained from hypothetical experiments, questionnaires and surveys (Johnsen et al. 2017 provide a broad overview). In general, previous studies have focused on the technical improvements (i.e., Naranjo & González, 2008; Lee, et al. 2015; Li et al. 2017), or their inclusion to the traffic (i.e., Fernandes & Nunes, 2012; Gold et al. 2016) and transportation (i.e., Crayton & Meier, 2017; Smith, 2013), as well as
on safety issues (i.e., Schoettle & Sivak, 2014; Fagnant & Kockelmann, 2015; Observatorio Cetelem Auto, 2016; Piao et al. 2016; Bock et al. 2017). Up to now, a common assumption is that sometime in the near future, AVs will have full autonomy on driving decisions, that is, with minimum or null human intervention. However, the human factor is still critical (Kalra & Groves, 2017). In part, this is because humans are responsible of the design of the AVs decision-making algorithms. Estimations reveal that more than 90% of the car accidents nowadays are provoked by human errors, and among the most relevant factors are those from their decision-making (NHTSA, 2008). Thus, to what extent decision related errors derived from behavioral considerations might be reflected in the AVs driving algorithms? In this respect, studies have shown that individuals who operate automated systems tend to exhibit gaps and mental misconceptions (e.g., Parasuraman et al. 2008; and Endsley, 2017). Furthermore, ethical and moral dilemmas add importantly to the aforementioned complexity. Solutions, until now, are limited. The well-established Software Engineering Code of Ethics documents the ethical and professional obligations of software engineers around the automation topic. The aim of the code is summarized in the following statement (see http://www.acm.org/about/se-code): “These principles should influence software engineers to consider broadly who is affected by their work; to examine if they and their colleagues are treating other human beings with due respect; to consider how the public, if reasonably well informed, would view their decisions; to analyze how the least empowered will be affected by their decisions; and to consider whether their acts would be judged worthy of the ideal professional working as a software engineer. In all these judgments concern for the health, safety and welfare of the public is primary; that is, the ‘Public Interest’ is central to this Code.”

Although research on the ethical dilemmas that AVs might face is still scarce, a set of hypothetical experiments have been conducted (see Holstein & Dodig-Crnkovic, 2018 for a comprehensive overview). In a seminal paper by Bonnefon et al. (2016) based on six on-line surveys, participants solved the Trolley Problem in different traffic contexts. They find that subjects preferred that AV algorithms sacrifice their passengers for a greater good. But paradoxically, they state that when riding an AV themselves, its algorithm should protect passengers at all costs. The study also reveals that individuals are less willing to buy an AV with a utilitarian decision-making algorithm. In other terms, although most of the interviewed subjects were supportive of a utilitarian ethics for the AVs, they also confessed their reluctance to travel in utilitarian AVs, knowing that their own lives could be sacrificed for the sake of the greater good. They supported utilitarianism for other people, but not for themselves. In line with this, Victor et al. (2018) provide experimental evidence suggesting
that other mental models, such as expectation mismatches, play an important role in the interaction human-automated vehicles.

It is also worth mentioning a recent paper by Awad et al. (2018) describing the results of the so-called Moral Machine, a worldwide survey that gathered almost 40 million responses to several moral dilemmas faced by hypothetical AVs. The researchers found strong evidence in favor of global preferences supporting sparing humans over animals, sparing more lives and sparing young lives, although they also found a great degree of variation in the responses due to cultural, demographic and economic reasons.

In an attempt to contribute to the debate about the role of cognitive models, the aim of this paper is to explore the effect of some important variables that were not taken into consideration in the literature. In the typical experiment, the subjects were only allowed to take basic decisions such as to swerve or not to swerve, and important features of the real world problems were set aside. For example, the consequences of our real world decisions are always surrounded by uncertainty, and it is usually difficult to predict with total accuracy what will happen after one decision has been taken. We want to take this uncertainty into account, as well as the possibility of taking some additional decisions that may influence such uncertainty, such as fixing the speed of the AV prior to a possible accident. The speed preferences are highly relevant in this context. It is considered a key factor with respect to driving preferences, subjective judgment of safety and risk, time saving, and speed enforcement (Tarko, 2012). Furthermore, the severity of injuries during a crash are closely related to the driving speed. Pedestrians, for example, have a 90% chance of survival when struck by a car travelling at 30 km/h, but the chances lower more than 50% when the speed is at 45 km/h (Pasanen, 1991). It is clear, then, that the speed of the AV has an influence on both the cost of the accident (the number of casualties) and its probability. Speed, cost and probabilities will be the basic variables in our experiment.

We also introduce other behavioral factors such as the aversion towards risk and loss, until now not considered in the related literature. In this respect, behavioral economists have concluded that human decision-making is so often irrational that their irrationality is predictable. What is then the rationality or the ‘predictable irrationality’ behind an ethical problem, in particular, when there are important economic components at stake?
A final issue that we study is related to the reliability of the laboratory experiments: to which extent the answers to the surveys and questionnaires can be used to predict the behavior of the subjects in the real world? This important issue has been studied in other contexts (see, for example, Gneezy & List (2006)), but to our knowledge, it has never been raised in relation to the moral preferences of the drivers. To this end, we have performed two types of experiments. In addition to the standard experiment where the subjects are asked about how the behaviour of a hypothetical AV should be programmed in a computer (we suggest to call this experimental design ‘ex situ’), we have also performed real driving simulations with a different set of subjects (we suggest to call this experimental design ‘in situ’). We have found significant differences in the results of both experimental designs, with interesting potential implications.

Besides our study, only Dickinson & Masclet (2018) have studied AVs and moral preferences from an economic experimental approach. As in our paper, they conduct a lab experiment to study choices from the Trolley problem. Yet, the experiments differ importantly in the following aspects.

- First, they adopt two versions of the ‘original Trolley dilemma’ while we (as well as Bonnefon et al.) work on the Trolley problem embedded in a drivers-decision-context (i.e. an “applied Trolley problem”).
- Second, their experiment employs a strategy elicitation method and a within-subject design where the Trolley problem was hypothetical and administered before a real payoff money-burning game. In comparison, we employ a direct elicitation method and use the classic between-subject design on a fully incentivized applied Trolley problem.
- And last but not least, we have performed two types of experiments, ‘in situ’ and ‘ex situ’, a design that allows us to compare the subjects’ behaviour when they actually drive a vehicle simulator versus when they program a hypothetical AV in the laboratory.

The outline of the paper is as follows: in Section 2, we present our behavioural predictions, the experimental design of our experiments is shown in Section 3 and in Appendix A, whereas the main results are commented in Section 4 and summarized in the Conclusions. The detailed instructions received by the participants in our experiments are presented in the Appendix B.

---

1 In both settings, the driving speed plays a role in the decision problem: in the in situ case the decision on the driving speed is as it occurs in a real setting, i.e., pressing the accelerator pedal, while in the ex situ this decision is programmed in the AV. With this setup, we are able to study the effect of speed in both conditions.

2 These two versions are in turn a reframed version of the seminal design by Foot, 1967
2. Behavioral Predictions

Policy-makers and automotive manufacturers assume that AVs will significantly reduce casualties by reducing the well-known human driving decision-making errors (NHTSA, 2008). However, the AVs standard behavior, up to know, is predetermined i.e., it is programmed by humans. We therefore should ask, how will people program AVs? Are we more or less selfish when we drive our vehicle versus when we program an AV? Can our actual driving behavior predict how humans will program the AVs decision-making algorithms? What is the role of the underlined economic consequences (including penalizations)? By using a controlled experiment, we aim to answer these and other related questions (see a detailed experimental design in Section 3).

2.1 The homoeconomicus driver

To simplify the terminology, from now on the “driver” will be either the person who actually drives a traditional car, or the person who programs an AV.

Motivated on the aforementioned questions, we are interested in eliciting actual choices\(^4\), and thus we present a fully incentivized experiment, meaning that each of the decision’s outcomes involves a real monetary payoff. Therefore, we first present our hypotheses (H) and predictions for the driver’s behaviour. Based on the standard economic theory i.e., a) individuals always maximize utility in any driving situation, and b) self-interest is the people’s primary motivation, we derive our main hypotheses:

**H1**: Drivers stay in their current path and hit the pedestrian

We also assume that drivers have stable preferences, therefore we have:

**H2**: Drivers exhibit the same driving preferences regardless of whether they actually drive or program an AV

Furthermore, in our experiment the monetary payoffs assigned to each of the driving choices, including the desired speed, have equal expected values. Thus from Expected Utility Theory (EU) it follows that:

---

\(^4\) Economists argue that incentives matter when eliciting preferences compared to eliciting preferences through hypothetical methods. However, we acknowledge a potential issue in using monetary incentives in a moral dilemma: the possibility of moral crowding out. Experiments manipulating the use of incentives might solve this issue and bring new relevant insights.
**H3:** Drivers are indifferent between driving at high or low speed

Yet, we consider that people typically give an important value to time i.e., people try to save time by doing things more quickly. Thus, if drivers impose a higher weight to gaining time (by driving faster), we should see a higher proportion of high-speed drivers.

**H3.1:** Drivers prefer a high driving speed

Although the standard economic model predicts unemotional and utility maximization choices, over a long period of time, influential economists including Smith (1759), Becker (1974), Arrow (1981), North (1990), Samuelson (1993) and Sen (1995), already pointed out that individuals do not always act in accordance with their material self-interest and often do care for the well-being of others. Moreover, ample empirical evidence supports these views (see for ex. Kahneman et al. 1982; De Bondt & Thaler, 1985; Camerer et al. 1997; O’Donoghue & Rabin, 1999; and extensive evidence thereafter).

### 2.2 Alternative Behavioral Assumptions

Behavioral concepts such as risk and loss aversion, moral beliefs, and bounded selfishness can explain potential deviations from our predictions above. In particular, we consider the role of other-regarding preferences i.e., those concerns for the well-being of others and desires to uphold social norms. In fact, such preferences have been shown to reduce social inefficiency in the absence of complete contracts (Arrow, 1971; Becker, 1976; Akerlof, 1984) and thus are key to solve social dilemmas (Ostrom, 1990). Thus, in this subsection we present our alternative hypotheses (aH).

If subjects exhibit pro-social preferences rather than egoistic we should expect that:

**aH4:** Drivers swerve to the side of the road and impact themselves against a barrier, therefore leaving the pedestrian completely unaffected

Yet, how does the exposure to risk affect choices? We present the Trolley problem as a typical binary gamble choice involving a sure option (stay on the current path) and another (swerve) with two possible payoffs (with equal expectations). Thus, if decision-making is driven by risk/loss aversion we have:

**aH5:** Drivers prefer the sure option (impact a pedestrian)
Moreover, given the well-known relationship between speeding and risk-attitudes on the roads (e.g., Oppenlander, 1966; Wasielwski, 1984; and evidence thereafter) we have:

aH6: Risk-averse drivers prefer a lower speed.


In our experiment, we let subjects solve a reframed version of the well-known Trolley Problem. In this dilemma, a driver of a vehicle is traveling down a main road and suddenly a pedestrian appears directly in the car’s path. The driver then has to decide whether to swerve to avoid an impact against the pedestrian, or to stay on the path and therefore hit the person. Using a between-subject design we compare decisions from two conditions: when subjects drive a vehicle simulator (SIM) and when subjects program in a PC the behavior of an hypothetical autonomous vehicle (AV) ⁶. Our reframed version of the Trolley Problem is as follows (detailed instructions in Appendix B).

In both conditions, subjects are informed that they were matched randomly and anonymously with another participant (Player X). They also received information about the gender of their match and that they, and their corresponding match, have been endowed with 6 euros each. Next, they learnt from the instructions that they represent the driver who is traveling down a main road. They also learnt that at some point during their trip (they do not know exactly when) a pedestrian avatar, who represents Player X, will suddenly appear ahead in the car’s path. Then they decide between swerving to avoid an impact against the pedestrian, or staying on the path and therefore hit Player X. In both conditions, participants were shown a pictorial representation of the decision situation similar to the one used in the Study 2 from Bonnefon et al. (see our template in Appendix A). They also learnt that their decision has real monetary consequences for them (driver), and for their match (pedestrian). Finally, they learnt that the severity of these consequences depends on their driving speed. Participants played both roles, the role of driver (active role) and the role of pedestrian (passive role), however, subjects were not informed about their passive role. Instead, they were informed at the end of the experiment that they also played the role of Player X and about the corresponding monetary outcomes. In SIM, the simulation of the Trolley Problem was programmed using an ad-hoc software from the simulator provider (SimuTech, Gesellschaft für Fahrsimulation mbH). Subjects were asked to get familiar with the simulator hardware (steering wheel, pedals, speed, screens, etc.) by making a driving test before the actual experiment. In a different way from

⁶ See pictures of the simulator device and the experimental setup in Appendix A.
the AV condition, in SIM, the simulator software estimates the average driving speed per driver. This is registered from the moment each subject starts the ride, until just the moment of impact, either against a barrier, or against the pedestrian. The ride finishes in the moment of the impact and has a maximum duration time of approx. 1 min. The vehicle speed ranges from 0 km/h up to a maximum of 220 km/h.

In the AV setting, subjects face the same Trolley dilemma as in SIM. In this case, however, instead of driving a vehicle, subjects solve the dilemma by programming in a PC a hypothetical autonomous vehicle (see pictures of the setup and detailed instructions in the Appendix). They also decide the driving speed by programming it at the same time they program their decision on whether to swerve or stay on the path. We decided this setup in AV for two reasons, to study the role of speed on driving choices in both treatments, and to make a fair comparison between them i.e., to avoid differences in decisions between treatments being attributed to the absence of the speed choice in AV. For simplicity, we will call the average driving speed in SIM and the programmed driving speed in AV, simply “speed”.

The economic outcomes for the driver and Player X in SIM and AV are as follows (see detailed instructions in Appendix B):

1. If the driver [programs the car to] swerves to the side of the road, the car will impact a barrier with two possible outcomes:
   
   a) the driver loses his 6 euros’ endowment or,

   b) the impact leaves the driver completely unaffected, therefore she keeps her money.

   - The probability that each outcome occurs depends on the driving speed.

      - If the driving is at the typical driving speed limit\(^7\), or higher (≥ 50km/h) then:

        There is a 50% chance that a) occurs, and a 50% chance that b) occurs.

      - If the driving is at a lower speed than the legally permitted (<50 km/h) then:

        There is a 25% chance that a) occurs, and a 75% chance that b) occurs.

\(^7\) We consider the typical driving speed limit within European urban areas as 50 km/h. During the ride in SIM, the driver is able to see several speed limit signs on the road, just as it occurs in real-driving situations.
Importantly, if the driver [programs the car to] swerves to the side of the road, the pedestrian is completely unaffected and therefore Player X keeps her own 6 euros endowment.

2. If the car stays on its current path, it will hit the pedestrian and therefore Player X loses her 6 euro endowment.
   - Here, the driver will also lose an amount depending on the driving speed.
     - If the driving is at speed limit or higher (≥ 50km/h), the driver loses 3 euros.
     - If the driving is at low speed (<50 km/h), the driver loses 1.5 euros.

Regardless of the decision taken 1) or 2), the driver receives an additional 1.5 euro if the driving speed is ≥ 50 km/h. The aim of this premium for driving faster is to mimic a monetary reward for saving time (monetary time value).

Notice that the expected values derived from any of the decisions taken are identical. Based on Expected Utility Theory, there are no economic reasons (except from the experimental monetary reward8), to induce any particular preference over the available options. Yet, setting the expected values of the available options to be the same does not mean they appear to subjects as equivalent. We might expect that the risk attitudes of the participants are different. To ensure full understanding of the instructions and payoffs, we asked participants to take their time to read the instructions. Once they stated that everything was clear, we asked them to solve several hypothetical exercises about the possible payoffs prior to the actual experiment. In these exercises, participants had to calculate both their own and Player X payoffs. The answers were reviewed before proceeding. When an answer was incorrect, a pop-up message with the correct answer appeared in the screen. In addition, subjects were told that they could raise their hand for further assistance.

In total, 150 students from the Neu-Ulm University of Applied Sciences, Germany, participated in the experiment: 71 in SIM and 79 in AV. In AV, participants were distributed over 5 sessions in one of the university’s PC-labs while in SIM, subjects participated in individual sessions using a driving-simulator and conducted at the simulator-lab. In both conditions, each session lasted about 40 minutes and participants were paid privately and in cash after the session was over. Participants earned on average 14 Euro including a 5 Euro compensation for their attendance to the experiment.

---

8 The reward medium is a key concept from the Induced Value Theory and arguably one of the methodological innovations for experimental economics. It is based on the idea that a suitable reward can induce pre-specified characteristics in the subjects so that other characteristics become irrelevant (Smith, 1976).
and additional incentives for answering a set of follow-up questions. This amount is relatively higher than the average payment that a student would earn in one typical working hour in Germany.

Subjects were in their early twenties (\textit{Mdn} = 23 years, \textit{SD} = 3.23), had mainly a business-economics background, and 50\% were female. Participants received detailed instructions through the PC screen and a summary of them on paper. The experiment was programmed using the experimental software z-Tree (Fischbacher, 2007). All sessions were conducted in German language.

3. Results

In this section, we summarize the main results. The descriptive statistics from Table 1 show that, on average, subjects solve the Trolley problem differently when they drive a simulator and when they program an AV. Overall, we see the following: for the SIM experiment, the proportion of subjects who impact Player X (\textit{Impact X}) at a high speed (.451) doubled the proportion of those who made the same decision at a low speed (.239); on the other side, the proportion of subjects who decided to program their AV to \textit{Impact X} at a high speed (.329), tripled the proportion of those who did it at a low speed (.114). The highest proportion goes to people selecting high speed SIM, whereas the lowest corresponds to those selecting low speed in AV.

<table>
<thead>
<tr>
<th>Table 1. Proportion of Decisions and Statistical Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIM</td>
</tr>
<tr>
<td>\textit{Impact Yourself}</td>
</tr>
<tr>
<td>High Speed</td>
</tr>
<tr>
<td>Low Speed</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Notes. \(^a\) Two-sample test of proportions (Acock, 2014)

Number of observations in (). In SIM \(n= 71\) and in AV \(n= 79\)

3.1 Decisions on sacrifice oneself vs sacrifice others

In order to determine whether the observed differences between SIM and AV are statistically significant, we compute the corresponding Two-sample test for difference in proportions. First, we see that the proportion of subjects who impact themselves (\textit{Impact Y}) and \textit{Impact X} (0.309 and .690) is significantly different in SIM compared with AV (.557 and .443), \(Z = 3.04, p < .01\). Overall, there is a higher proportion who \textit{Impact Y} in AV vs. SIM. Then, on average, the standard economic
model (Hypothesis H1) correctly predicted behavior in SIM, but its predictions failed in AV, which calls for alternative theoretical models (Hypothesis aH4. See also our regression models in Section 3.6). Moreover, driving preferences are not stable across treatments, thus Hypothesis H2 is rejected. Choices in the AV condition then, contrast with survey responses in Bonnefon et al. where participants strongly disagreed that it would be more moral for AVs to sacrifice their own passenger to save the life of one pedestrian. They find that 23% of their participants think that AVs should sacrifice their passengers when only one pedestrian could be saved (Study 2). This percentage increased with the number of pedestrians that could be saved. The same tendency was observed when subjects were asked their preference toward a legally enforced utilitarian sacrifice (Study 5). Also, our findings differ from previous studies showing that subjects prefer no action (in our setup it represents to stay in the current path-act of omission) over action (swerve-act of commission). This behavior is based on the “action principle” which indicates that in the eyes of individuals and courts of law, an act of omission is more morally accepted than an act of commission that results in similar consequences (see Spranca et al. 1991; Cox et al. 2017 and Dickinson & Masclet, 2018). What is then the role of morality in driving decisions? We formally address this question later in this section.

3.2 Decisions on the speed

Now, we turn to the analysis of the speed preferences. In line with Hypothesis H3.1, but in contrast with Hypothesis H3, we find a higher proportion of subjects who prefer a high speed rather than a low speed (.58 vs .42). The same tendency holds within treatments (.59 vs .41 in SIM and .57 vs .43 in AV). However, for a deeper analysis, in Table 1 we divide the chosen level of speed by steering choices (Impact Y vs Impact X) and by treatment. According to the first raw, we see that the proportions of steering decisions among high-speed subjects in SIM (.141 and .451) and in AV (.241 and .329) are statistically different at p<.10. This is also true for low-speeders at p= .010 significance level (.169 and .239 in SIM and .316 and .114 in AV). We also see that in AV, the proportion of subjects who decide to impact themselves at high speed (.241) is significantly higher than the proportion in SIM (.141). Note that similar behavior is observed among the low-speed group. Here the proportion of drivers who impact themselves in AV (.316) is also significantly higher than the proportion in SIM (.169). These tendencies might indicate that decisions to solve the Trolley Problem while driving a simulator or programming an AV are, on average, not conditional on the speed choices. We see, however, that the differences in proportions are stronger
among low-speeders compared to high speeders (p=.010 vs p=.069). Later on, we formally test these results by running a set of regression models.

### 3.3 Analysis of choices within each experimental condition.

Finally, we compare choices within each experimental condition. First, we see that in SIM, the proportion of subjects who *Impact Y* (.310) is significantly lower than the proportion who *Impact X* (.690). The opposite is true in the AV condition. Here the proportion who *Impact Y* (.557) is higher than the proportion who *Impact X* (.443).

Regarding the choices on speed, we see that in SIM, the proportions of subjects who impact themselves and impact another participant at high and low speed are not statistically different ($Z=1.57; p=.116$). We see, however, that the proportion of subjects who *Impact X* at high speed (.451) doubles the proportion of those who made the same decision at low speed (.239). We do not observe a significant difference between high-low speeders who *Impact Y* (.141 vs .169). Decisions in AV follow a similar tendency. Here, however, the difference in choices at high and low speed are highly significant ($Z=2.77; p<.01$). More specifically, we see that the proportion of subjects who decided to program their AV to *Impact X* at high speed (.329), triples the proportion who did it at low speed (.114). In sum, results on speed preferences suggest that speed by itself is not a significant factor for driving choices between SIM and AV. Instead, speed seems to be particularly relevant when subjects decide to *Impact X*, independently on the experimental condition.

### 3.4 Probit regression models

To statistically confirm our results and further investigate the driving behavior, we present in Tables 2-4 a set of Probit and Multinomial regression models. Overall, regressions confirm that there are significant differences between solving the Trolley Problem in SIM vs in AV.

From Table 2 we see that the negative coefficient for *Treatment* in regression (1) confirms that programming an AV significantly reduces the probability of impacting another person (but increases the chances of impacting themselves). This is robust after controlling for Speed (2), gender, and the gender of the pedestrian $^{10}$ (3). In contrast with a gender effect found in Bracht & Zylbersztejn (2017) and in Dickinson & Masclet (2018), our coefficient for *Female* and *Female_O*

---

$^{9}$ As a robust test, we run in R a set of Logit regressions. We obtained similar results to those presented in this paper.  
$^{10}$ New variables, such as the number of passengers and pedestrians or their socio-demographic characteristics, might provide relevant relationships with driving behavior and with programming AVs preferences.
indicates that decisions in our experiment are not related to the driver’s gender or on the pedestrian’s\textsuperscript{11}. In addition, our findings contribute to the risk assessment that insurance firms have historically conducted on gender (Williams, 1992). We must state that this difference between our results and previous findings can be attributed to the fact that our design involves only one pedestrian, while the aforementioned studies deal with utilitarian choices i.e., the number of potential victims is higher. Next, the negative coefficient for Speed in (2) and (3) shows that low-speed drivers have less propensity to impact a pedestrian than high-speed drivers, while the latter are less willing to sacrifice themselves for the good of another participant. Given the positive association between speeding and risk-taking described in the Hypothesis aH6, these results might indicate that low-speed (risk-averse) drivers exhibit more pro-social behavior compared to high-speed (risk-seeking) drivers who seem to be more self-centered\textsuperscript{12}. In line with this hypothesis, Machin & Sankey (2008) report that personality traits such as risk-aversion strongly correlate with speeding behavior among young drivers. Interestingly, they found that lower levels of altruism (which is defined as a concern for the welfare of others i.e., prosocial behavior) are also associated with speeding, while Vassallo et al. (2007), find that young drivers who were more antisocial reported also greater speeding. Later on in this section, we show that driving choices in our experiment significantly correlate with the drivers’ risk-aversion. Could it be that these speed preferences change depending on whether one drives a traditional car versus an autonomous vehicle? The coefficient for Treatment$\times$Speed in (4) confirms that the speed effect applies for both settings.

To further investigate the speed choices, we now consider Speed as the dependent variable. From Table 3 we see that the coefficients for Treatment confirm that speed preferences are not dependent on whether programming an AV or actually driving. Moreover, the coefficient for Impact X also confirms that low-speed drivers are less likely to impact a pedestrian. Regarding gender effects, we see in (3) that only the coefficient for Female\textunderscore O is significant. However, this effect disappears when we include in (4) the interaction with Impact X.

Now, given that the steering and speed choices are made simultaneously, we examine these two decisions as a single choice i.e., we construct four categories: Impact Y at high/low speed, and

\textsuperscript{11} We controlled for the fact that both players had the same gender (about half (.49) of the matches had the same gender). We find that same-gender has no effect on driving choices with $p=.224$. The same result is true when we consider an interaction with Treatment with $p=.918$.

\textsuperscript{12} We do not find statistical evidence of correlation between our measures of risk-aversion and speed.
Impact X at high/low speed. Thus, we present in Table 3 a set of Multinomial Probit models (1-4). Supporting our aforementioned findings, we see in (a) that subjects who program an AV are less (more) likely to impact Player X (themselves) either at high or low speed, while Treatment has no significant effect on (2). The treatment effect holds after controlling for the driver’s (Female) and Player X’s (Female_O) gender. Finally, Female_O suggests a slightly significant effect (p<.10) in (1) and (2), however, this slight effect vanishes in (3).

3.5 Driving choices and their relationship with behavioral covariates

After subjects solved the Trolley dilemma, and before they received the outcomes of the experiment, participants answered a set of incentivized follow-up questions. Subjects received on average 5 euros for solving the following well-established elicitation methods. First, we elicited risk and loss aversion using the decision problems by Kahneman & Tverksy, (1984, 1992). For the former, participants decided between these two options\textsuperscript{13}:

\begin{itemize}
  \item [a)] "Receive 240 euros for sure" or
  \item [b)] "Play a gamble with 25% of receiving 1000 euros, and 75% of receiving nothing".
\end{itemize}

For the latter, subjects stated the amount (x) that they would want to win in order to be indifferent between these two options\textsuperscript{14}:

\begin{itemize}
  \item [a)] “Receive 0 euros for sure” or
  \item [b)] “Flip a coin where you have 50% probability of receiving x, and 50% probability of losing 25 euros?”
\end{itemize}

We choose these procedures due to their simplicity and clarity. Subjects can easily understand the task, make the calculations of the expected payoffs, and identify the difference between the options (risk and loss)\textsuperscript{15}. Finally, we elicited the participants’ moral beliefs about how the Trolley dilemma should be solved. Here, we used a reframed version of the question presented in Bonnefon et al. (2016)\textsuperscript{16}.

\textsuperscript{13} 123 subjects (=83\%) revealed risk-aversion by choosing option a) 
\textsuperscript{14} Loss neutrality = 25, M= 246, Mdn= 100 
\textsuperscript{15} We muss stress that other robust methods can provide more refined estimations. 
\textsuperscript{16} We find a M= 62 and Mdn= 59 while in Bonnefon et al. Mdn= 85. We are aware that other elicitation methods might bring new insights on driving behavior.
“From a scale where 0= protect the passenger and 100= protect the pedestrian, rate what is the most moral way to drive [program an AV].”

3.6 Behavioral determinants of driving choices.

In models (5) and (6) from Table 2 we analyze the effect of behavioral factors on driving behavior. In contrast with our Hypothesis aH5, the negative coefficient for Risk Aversion in (5) reveals that, individuals who are typically reluctant to take risks, have more propensity to protect pedestrians by sacrificing themselves. Moreover, from the interaction term in (6), we see that risk-aversion has a larger effect when individuals program an AV than when they actually drive. This runs counter to common intuition, since we would expect a risk-averse driver to continue in the current path (at low speed) and impact a pedestrian because with this decision she avoids the 50-50 gamble represented by crashing himself against a barrier. We then provide the following explanation: risk-averse individuals on the road seem to exhibit more social preferences compared to risk-lovers. This might be because under extreme traffic situations (as the one presented in this experiment), risk-aversion seems to be offset by other-regarding preferences, a mechanism that seems not to operate among risk-seeking drivers who exhibit a more selfish behavior. This is in line with our results on speed choices. Preferences for low-speed driving (typically preferences related to risk-aversion) are also correlated with the propensity to protect pedestrians.

We find no statistical evidence that Loss-Aversion and beliefs about Morality play a role on the Trolley problem decisions. Finally, when we treat Speed as the dependent variable, we find that our behavioral covariates are not correlated with Speed choices.
Table 2.

Determinants of the Trolley Dilemma Decisions

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>-.639*** (.211)</td>
<td>-.658*** (.216)</td>
<td>-.611*** (.222)</td>
<td>-.449 (.290)</td>
<td>-.665*** (.227)</td>
<td>.192 (.566)</td>
</tr>
<tr>
<td>Speed</td>
<td>-.670*** (.216)</td>
<td>-.718*** (.222)</td>
<td>-.514 (.323)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-.236 (.219)</td>
<td>-.235 (.219)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female_O</td>
<td>.186 (.221)</td>
<td>.211 (.224)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment*Speed</td>
<td></td>
<td></td>
<td></td>
<td>-.385 (.445)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Risk Aversion</td>
<td></td>
<td>-.658** (.314)</td>
<td>-.054 (.452)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss Aversion</td>
<td>-.000 (.000)</td>
<td>-.000 (.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morality</td>
<td>.000 (.004)</td>
<td>.001 (.004)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment*Risk Aversion</td>
<td>-.994* (.600)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>.496*** (.156)</td>
<td>.792*** (.191)</td>
<td>.801*** (.229)</td>
<td>.697*** (.252)</td>
<td>1.10*** (.426)</td>
<td>.497 (.537)</td>
</tr>
<tr>
<td>P &gt; Chi²</td>
<td>.045</td>
<td>.093</td>
<td>.097</td>
<td>.101</td>
<td>.071</td>
<td>.084</td>
</tr>
</tbody>
</table>

Notes. Probit regression models ( ) with robust std errors in parentheses.
Dichotomous Dependent Variable where 0= impact yourself and 1= impact player X
Treatment =1 for AV. Speed =1 when speed is < than 50km/h (speed limit).
Female and Female_O = 1 when passenger and Player X is female respectively.
Risk Aversion =1. Loss Aversion: continuous covariate (loss neutrality = 25, M = 246 Mdn = 100)
Morality: continuous covariate from a scale where 0= protect passenger and 100= protect pedestrian (M = 62 Mdn = 59)
n = 150 for all models.
***p < .01, **p < .05, *p < .10
<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>.056</td>
<td>-.118</td>
<td>-.166</td>
<td>-.163</td>
</tr>
<tr>
<td></td>
<td>(.207)</td>
<td>(.218)</td>
<td>(.225)</td>
<td>(.227)</td>
</tr>
<tr>
<td>Impact X</td>
<td>-.680***</td>
<td>-.724***</td>
<td>-.756**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.219)</td>
<td>(.223)</td>
<td>(.302)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-.114</td>
<td>-.119</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.219)</td>
<td>(.220)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female_O</td>
<td>.458**</td>
<td>.422</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.216)</td>
<td>(.321)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female_O*Impact X</td>
<td>.066</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.439)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-.231</td>
<td>.231</td>
<td>.115</td>
<td>.132</td>
</tr>
<tr>
<td></td>
<td>(.150)</td>
<td>(.214)</td>
<td>(.257)</td>
<td>(.275)</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>.005</td>
<td>' &lt; .001</td>
<td>.048</td>
<td>.073</td>
</tr>
<tr>
<td>P &gt; Chi²</td>
<td>.786</td>
<td>.007</td>
<td>.007</td>
<td>.014</td>
</tr>
</tbody>
</table>

Notes. Probit regression models ( ) with robust std errors in parentheses.
Dichotomous Dependent Variable where 0= high speed and 1= low speed (< than 50km/h)
Treatment =1 for AV. Impact X =1 if impact Player X.
Female and Female_O = 1 if impact Player X.
n = 150 for all models.
***p < .01, **p < .05, *p < .10
Finally, for behavioral and risk-assessment studies it is important to measure the performance of the prediction method. Therefore, we estimate the area under the receiver operating characteristic (ROC) curve from our Probit regression model. We find that the AUC (area under the curve) takes the value .707 indicating a strong predictive validity of our model\(^\text{17}\) (see Fig. 1).

\[\text{Figure 1: Area under the ROC}\]

\[\text{Table 4. Multinomial Probit Models of the Troley Dilema Decisions}\]

<table>
<thead>
<tr>
<th>Impact Y-High Speed (1)</th>
<th>Treatment(^a)</th>
<th>Intercept(^a)</th>
<th>Treatment</th>
<th>Female</th>
<th>Female_O</th>
<th>Intercept</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.866**</td>
<td>-.364</td>
<td>.888**</td>
<td>.400</td>
<td>-.740*</td>
<td>-.183</td>
</tr>
<tr>
<td></td>
<td>(.381)</td>
<td>(.271)</td>
<td>(.395)</td>
<td>(.388)</td>
<td>(.388)</td>
<td>(.360)</td>
</tr>
<tr>
<td>Impact X-High Speed (2)</td>
<td>.243</td>
<td>.493**</td>
<td>.326</td>
<td>.153</td>
<td>-.631*</td>
<td>.705**</td>
</tr>
<tr>
<td></td>
<td>(.347)</td>
<td>(.231)</td>
<td>(.361)</td>
<td>(.358)</td>
<td>(.357)</td>
<td>(.325)</td>
</tr>
<tr>
<td>Impact Y-Low Speed (3)</td>
<td>.950***</td>
<td>-.243</td>
<td>.927**</td>
<td>.297</td>
<td>-.307</td>
<td>-.177</td>
</tr>
<tr>
<td></td>
<td>(.370)</td>
<td>(.263)</td>
<td>(.376)</td>
<td>(.375)</td>
<td>(.371)</td>
<td>(.368)</td>
</tr>
<tr>
<td>Impact X-Low Speed (4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Base Outcome</td>
</tr>
<tr>
<td>P &gt; Chi²</td>
<td>.019(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.099</td>
</tr>
</tbody>
</table>

Notes: Cateogrical Dependent Variables by driving choice (Impact Y/X) and speed choice (high /low). 
\(a = \) Model without controls. \(\text{Impact X-Low Speed} = \) Outcome Base 
Low speed < 50km/h (speed limit). \(\text{Treatment} = 1 \) for AV. 
\(\text{Female} \) and \(\text{Female}_O = 1 \) when passenger and Player X is female respectively. 
Models with robust std errors in (). 
\(n = 150 \) for all models. 
\(***p < .01, **p < .05, *p < .10\)

\(^{17}\) In applied psychology and prediction of future behavior, it is considered that AUC values of .70 and higher indicate strong predictive effects (see for example Rice & Harris, 2005).
4. Conclusion

The advent of driving assistance systems has raised high expectations about important benefits, not only for the automotive industry, but also for society. Even higher expectations have raised around fully autonomous driving, mainly in terms of safety. Considering the 1.4 million fatalities per year on worldwide roads (Statistics from the Global status report on road safety 2018), it is not surprising that AVs will help to solve this serious problem by minimizing the typical human-driving errors. However, we should not forget that AVs are human ‘creatures’, and that their behaviour is always predetermined, since it is the result of following some programming rules. In other words, the AV algorithms are designed by humans, and the decision-making involved might either reflect the preferences from a limited number of people, say the programmers, or from a higher number, a larger proportion of society. A key factor with respect to driving preferences is the speed. This is because it is closely related to our own and others’ safety, time saving, and speed law enforcement. More specifically, driving speed is related to the probability of an accident and to the severity of injuries and other economic costs. All these elements have motivated us to study how our future AVs will decide on behalf of us when we face risky traffic situations e.g. deciding between swerving and crashing our car against a barrier to avoid an impact against others, or continuing on our current path and therefore affect others. Using a reframed version of the well-known Trolley-problem, we investigate how individuals face this dilemma when they actually drive a vehicle simulator versus when they solve it by programming a hypothetical AV in the laboratory.

In general, it could be argued that there are three main ways of programming the behavior of an AV:

1) To use ‘a priori’ ethical rules, based on philosophical reasoning. The problem with this approach, however, is that there is not a clear consensus about what these ‘a priori’ rules should be.

2) To ask people about the appropriate rules to follow when facing an inevitable accident. This is what we call the ‘ex situ’ approach.

3) To infer the rules from the behavior of the real drivers, the ‘in situ’ approach.

Our experiment was carried out under the second and third procedures: our ‘in situ’ method is based on the subjects’ revealed preferences regarding hypothetical AV programming, while in the ‘ex situ’ case, it is based on the recorded behavior observed in a driving simulator. Furthermore, and in contrast with previous research, we consider the role of economic incentives, probabilities...
and magnitude of accidental outcomes as a function of driving speed, as well as behavioral factors such as moral beliefs, and risk and loss-aversion.

Our results show a significant difference between what people say ‘should be done’ ex situ and what they actually do in situ: individuals who program an AV and who are more cautious in terms of speed are less (more) likely to sacrifice a pedestrian (themselves) compared to those who actually drive and prefer a higher driving speed. Moreover, we find that these differences are associated with risk-aversion, and that this effect is particularly strong when subjects program an AV at low speed. We might venture that this is because risk-averse individuals exhibit stronger other-regarding preferences compared with risk-seekers.

Our conclusions are in line with those obtained in other research areas in decision-making. According to the Construal Level theory, subjects perceive distant events (e.g. programming an AV) as more morally important than proximal events (actual drive), and temporal distance can lead to harsher moral judgments (Agerström & Björklund, 2009), and greater pro-social intentions (Henderson et al. 2006 and Choi et al. 2012). Other economic experiments have detected that an increase of the deliberation time in decision problems can lead to an increase of cooperation: in other terms, that ‘cold’ decisions are more cooperative than ‘hot’ decisions. For example, Oechssler et al. (2015) find that a 24-hour cooling-off period leads to fewer rejections in ultimatum bargaining, and Brandts & Charness (2011) find the levels of punishment lower when the subjects make conditional responses for each possible situation (the ‘strategy method’ of eliciting preferences) instead of a direct-response method. In addition, the proximity and promptness of the real driving experiences involve emotional and unconscious reactions that can trigger self-preserving responses.

Finally, there is a growing debate on how and when the use of monetary incentives modify behavior (see Gneezy et al. 2011 for a comprehensive overview). In particular, less is known about the effects of (extrinsic) incentives to revert the high number of fatalities per year in worldwide roads. Evidence shows that speed is one of the main driving decisions and thus responsible, to some extent, for millions of accidents (e.g., West & Hall, 1997; Elvik, 2008; New South Wales Centre for Road Safety, 2008; NHTSA, 2008). Therefore, assigning incentives (penalizations) to lower (higher) driving speeds may help to change risky driving behaviors. Regardless of the experimental setting, our results uncovered a high proportion of subjects who were willing to sacrifice another participant even when this represented a monetary penalization, which in addition, increased when the driving speed was higher.
The question of which response is more policy-relevant remains: should the AVs be programmed based on observed real driving behavior or on answers to driving dilemmas in the laboratory? Whose preferences should be considered? Should our AV decision-making algorithms reflect the programmers’ preferences, or should they consider a larger part of our society? The current paper does not go this far since there are other factors that are still uncovered. Yet our results exhibit an important difference between implementing drivers’ behavior in automated driving algorithms by means of in situ machine learning on the one hand, and by means of coding ex situ choices on the other. We thus consider that this discrepancy should be further investigated, bearing in mind that the design of future AV driving algorithms is somehow a reflection of human preferences.
References.


Li, G., Li, S. E., Cheng, B., & Green, P. (2017). Estimation of driving style in naturalistic highway


Wasielewski, P. (1966). Speed as a measure of driver risk: observed speeds versus driver and vehicle characteristics. *Accident Analysis & Prevention, 16*(2), 89-103.


Appendix A Experimental setup

Figure A1: Driving simulator (Simu Tech GmbH, Bremen, 2017) of a traditional vehicle: automatic version, three screens for surrounding effects, speedometer (max. 220 km/h), three rearview mirrors, light indicators, ignition, gas and breaks pedals, and steering wheel.

Figure A2: Follow-up questions setup in SIM using z-Tree (Fischbacher, 2007).
Figure A3: Experimental setup in AV using z-Tree (Fischbacher, 2007).

Figure A4: Template representing the trolley problem shown in the instructions for both SIM and AV.
Appendix B Instructions.

You have received 6 euros to participate in the experiment. You have been matched with Player X who is another participant located in this room, however, you will not know with whom you have been matched neither during nor after the experiment.

You are the sole passenger in a vehicle [an autonomous vehicle] traveling down a main road. Suddenly, Player X appears ahead, in direct path of the car!

Now you have to decide whether [The car can be programmed] to swerve to the side of the road, or stay on the current path.

1. If you [the car] swerve to the side of the road, you will impact a barrier with two possible outcomes: a) You lose your 6 euros or,

   b) The impact leaves you completely unaffected, then you keep your 6 euros.

   The probability that each outcome occurs depends on the speed you [program the car] drive.

   - If you [program the car to] drive at speed limit or higher (≥ 50km/h):
     
     There is a 50% chance that a) occurs and a 50% chance that b) occurs.

   - If you [program the car to] drive at a lower speed (<50 km / h):
     
     There is a 25% chance that a) occurs and a 75% chance that b) occurs.

A random device will decide one of the two possible outcomes a) or b).

Notice that if you [program the car to] swerve to the side of the road, Player X is completely unaffected. Then he/she keeps his/her 6 Euro endowment.

2. If you [program the car to] stay on the current path, you will impact Player X making him/her lose his/her 6 Euros endowment.

   Notice that if you [program the car to] stay on the current path, you will lose and amount of money depending on the speed you drive.

   - If you [program the car to] drive at speed limit or higher (≥ 50km/h):
     
     You lose your 6 euro.
- If you [program the car to] drive at a lower speed (<50 km / h):

You lose 3 euro.

Information about the speed:

Since “time is money”, you will receive an additional 1.5 euros if you [program the car to] drive at speed limit or higher no matter whether you [program the car to] SWERVE or STAY on the current path.