ETHICS OR SELF PRESERVATION? AN ONLINE STUDY EXAMINING DRIVER RESPONSE TO ON-ROAD OBSTACLES DURING AUTOMATED DRIVING

Helena Kaul

Wichita State University
1845 Fairmount St.
Wichita, KS 67260

ABSTRACT

In the trolley problem paradigm, a person is faced with an ethical dilemma where they must decide how to distribute inevitable loss of life such as deciding between letting five people die on the tracks in front of a trolley or pulling a lever that causes the trolley to switch to a separate track and kill one person. This online study asked participants to monitor a simulated automated vehicle and intervene if they felt the vehicle should change lanes. The results found that participants intervened roughly 96% of the time when the group of five bollards was in front of them, whether this caused them to enter an empty lane or a lane with a single alternative bollard. This suggests that drivers may respond randomly when forced to make a decision under pressure, which can lead to a worse situation.

1. INTRODUCTION

The trolley problem is a well-known thought experiment where a person must decide between letting five people die on the tracks in front of a trolley or pulling a lever that causes the trolley to switch to a separate track and kill one person instead. Past studies have found that, in trolley problem scenarios, people tend to prefer the utilitarian decision that saves the most lives (Navarrete et al., 2012). A previous online study found that people reacted more randomly as Time-to-Collision (TTC) decreased to 1 second before the impact with pedestrians (Yahoodik et al., 2021), suggesting that people may need at least two seconds to generate an ethical decision that follows their ethical preference (e.g., Samuel et al., 2020). The purpose of this study is to examine whether the effect of TTC on participants' binary decision persists with bollard avatars rather than pedestrian avatars due to the less ethical scenario.

2. METHODS

Thirty participants were recruited from the community of Old Dominion University, Norfolk, VA. Eleven participants were excluded because they either failed a catch trial during the experiment or completed more than one experimental session. Most of the participants were licensed drivers from 0-7 years with only three people reporting that they had a license for more than 10 years (M = 5.39 years, SD = 5.67). The data from the remaining 19 participants were used for the analysis reported below.

The study employed a 2 x 2 x 3 within-subjects design with three factors: placement of bollards (in the right lane vs. the left lane), alternative bollard (present vs. absent in the opposite lane), and TTC (1, 2, vs. 3 seconds). The experiment was hosted in Pavlovia.org and conducted remotely. The participants viewed 16 videos rendered by a driving simulator in a random order. The experiment was conducted in PsychoPy. They were instructed to imagine that they were piloting a partially automated vehicle and to press the space bar on their keyboard if they wanted the car to switch to the opposite lane or withhold their response if they wanted the car to remain in the right lane. Four trials involved no

bollard avatars at all and served as catch trials. The data collected from any participants that pressed the space bar during the catch trials were excluded from the data analysis.

3. RESULTS

A logistic regression analysis was conducted using the generalized linear model function on R. There was a significant effect of Version, with participants being more likely to press the space bar when there was a group of bollards in front of them (96%) compared to when the bollards were in the other lane (54%), β = -6.144, p = .025, OR = 1/465.11. Participants pressed the space bar 54% of the time when the single alternative bollard was in front of them even though this would ultimately cause them to hit the group of five bollards in the other lane. Unlike in Yahoodik et al. (2021), we did not find a significant effect of TTC on response patterns. However, anecdotally, when a single bollard was in front of the participants, participants pressed the space bar to veer into the lane of the five bollards 89% of the time when the TTC was 1 second, compared to 36% of the time for the TTC of 2 and 3 seconds. This difference may suggest that participants react more randomly, deviating from the responses predicted by utilitarianism, when forced to make a decision under time pressure resulting in a decision that makes their situation worse. At the end of the study the participants were asked to rate their perceived acceptability of utilitarian ethics in automated driving on a scale from one (strongly agree) to five (strongly disagree). On average, participants moderately felt that utilitarian decision making was acceptable (M = 2.79, SD = 1.27).

4. DISCUSSION

A previous online experiment showed that participants responded consistent to the prediction of the utilitarian decision that many of them preferred, and their responses became more random as TTC decreased to 1 second before the impact with pedestrians. The current study examined whether this pattern persists even with unanimated objects such as bollards instead of pedestrian avatars using the online platform used in Yahoodik et al. (2021). The results showed responses like those reported in Yahoodik et al. (2021), that participants preferred to change lanes when they were approaching five bollards and remain when the alternative response leads to a collision with five bollards. This may suggest that people do not see much difference between pedestrian avatars and bollard avatars, but it also might mean that people avoid groups of obstacles out of self-preservation as well as not wanting to harm people. However, the current study showed no evidence for the effect of TTC on their response patterns. Anecdotal evidence, though, suggests that participants' responses are trending to be more random as TTC decreases. Further study with a sample size similar to that used in Yahoodik et al. (2021) would be necessary to examine whether driver responses vary as a function of TTC when imminent hazards are bollard but not human avatars. The current online experiment using videos rendered via a driving simulator only collected a binary response of a key press or the absence of a response each trial, without actually measuring steering or braking responses. Future research could analyze vehicle control performance and eye movements in a driving simulator.

REFERENCES

Navarrete, D. C., McDonald, M. M., Mott, Michael, L., & Asher, B. (2012). Virtual morality: Emotion and action in a simulated three-dimensional "trolley problem". Emotion, 12(2), 364-370.

Yahoodik, S., Samuel, S., Yamani, Y. (2021). Ethical decision making under time pressure: an online study. Proceedings of Human Factors and Ergonomics Society Annual Meeting, October 2021.

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, 1-7.