Driving on autopilot: The ethics of self-driving cars

Educator guide

PAPER DETAILS

Original title: The social dilemma of autonomous vehicles

Reference: Vol. 352, Issue 6293, pp. 1573-1576

Authors: Jean-Francois Bonnefon, Azim Shariff, Iyad Rahwan

Issue name: *Science*

Original publication date: 24 June 2016

DOI: 10.1126/science.aaf2654

Annotator(s): Alane Lim

TABLE OF CONTENTS

DISCUSSION QUESTIONS

1. What is the central philosophical question the authors address? How do they address it?	**SEP1** **SEP4** **SP2** **SP5** **SP7** **Nature of Science**
2. What are some of the problems with choosing to collect data using Amazon Mechanical Turk? How does this affect the way that the authors can generalize their results to the whole U.S. population?	**SEP4** **ETS2.B** **SP2** **SP7** **Data Collection**
3. What types of questions and survey techniques did the authors use to probe how participants felt about certain topics? What are the pros and cons of using surveys to collect this data (for example, relative to collecting data by making observations)?	**SEP4** **ETS2.B** **SP2** **SP7** **Data Collection**
4. Does the social dilemma in this paper apply to other technologies, as well? Can you think of any other philosophical dilemmas that might be introduced with a new technology?	**SEP1** **ETS2.B** **Cause and Effect** **SP7** **Nature of Science**
5. Based on the results of this study, what are some of the problems associated with government regulation of self-driving cars? What are some of the problems that might come up if the government does not impose regulations?	**ETS2.B** **Cause and Effect** **SP7** **Nature of Science**
6. Why do you think people are less willing to buy self-driving cars whose moral algorithms have been regulated by the government? Why do you think participants were less likely to think that autonomous vehicles would actually be programmed with a moral algorithm?	**ETS2.B** **Cause and Effect** **SP7** **Nature of Science**
7. The authors mentioned that they did not consider scenarios where the outcome of a crash is uncertain, or those that included information about the pedestrians such as age, gender, etc. Consider how adding this information would affect the morality of the decision. For example, is there a difference between sacrificing a young passenger to save an elderly pedestrian, or vice versa?	**SEP1** **SEP4** **ETS2.B** **Cause and Effect** **SP7** **Data Collection** **Nature of Science**

ACTIVITIES FOR INTERACTIVE ENGAGEMENT

Writing an abstract

Students write a new abstract for the article at a grade-appropriate reading level.

RST.9-10.2
RST.11-12.2
Nature of Science

Locating this study in the larger field

Students use the annotated list of references to explain how this research builds on the published work of at least one other independent group of scientists. Students will evaluate whether data from this research supports or contradicts previous conclusions, and reflect on the statement that scientific knowledge is a "community effort."

RST.9-10.8
RST.11-12.8
Nature of Science

Science in the news

Students explore news stories in the Related Resources tab and evaluate the stories for tone, accuracy, missing information, etc. Students may then write their own popular press article, at a grade appropriate level, about perception of driverless cars. They should be sure to address any weaknesses they identified in previously published pieces.

SEP4
RST.9-10.5
RST.9-10.6
RST.11-12.1
RST.11-12.5
SP7

Moral Machine at MIT

The Moral Machine (http://moralmachine.mit.edu/) is a simulation platform that the Massachusetts Institute of Technology uses to understand how people make moral decisions. Participants are presented with a series of moral dilemmas and must choose the lesser of two evils, such as killing two passengers or five pedestrians. The dilemmas presented in moral machine are similar to those used in this study, though they are more complex.

ETS2.B
Nature of Science

The next steps

Students design a follow-on experiment to this study that addresses either flaws or unanswered questions in the research at hand, or builds on it to explore a new question.

SEP1
RST.11-12.6
Nature of Science

ARTICLE OVERVIEW

Article summary (recommended for educator-use only)

Self-driving cars present an ethical dilemma. If self-driving cars are faced with the difficult decision of sacrificing either their passenger or pedestrians on the road, will they act in their passengers' interest, or will they sacrifice the passenger to save more lives? Through surveys, the researchers shed light on this question from a number of angles. Participants imagined a variety of scenarios and rated how they felt about various aspects of self-driving cars, including the morality of the car's sacrifice in different situations and whether they would want to buy a car whose moral decision making was legally regulated. Participants thought that cars which sacrificed their passengers for the greater good were highly moral, but indicated that they would not want to buy such cars themselves. Furthermore, participants did not want the government to regulate these moral algorithms in cars and would be less likely to buy a car subject to those regulations.

Importance of this research

This study was intended to provoke a discussion on the ethics of self-driving cars and the types of moral decisions people are willing to accept from self-driving cars. The questions the study asks are not necessarily novel, but the study shows how experimental ethics can shed insights on the complex moral, cultural, and legal issues that must be taken into account as self-driving cars become more ubiquitous. This has implications especially for carmakers and regulators, who will have to grapple with issues like who takes the blame if a self-driving car kills its passenger in a similar situation. The study notes that the situations in this work are even more complex in the real world. For example, the situations in the study did not really take into account the ages of the passenger(s) versus the pedestrian(s), nor did it consider uncertainty in the outcome of a collision (i.e., all collisions were considered to be fatal to at least one person). Future work can continue to build on these questions.

Experimental methods

- Online surveys: Participants imagined themselves in different scenarios, and answered questions in various ways, including answering whether a given course of action was moral, rating how likely they would do something (such as purchasing a vehicle) on a sliding scale from 0 to 100, and assigning points to different algorithms out of a budget of 100 points. The authors used Amazon Mechanical Turk to collect their data.
- Regression analysis: Regression analysis was done to determine potential confounding variables (covariates). Regressions were then performed on subsequent studies, obtaining p-values that showed the significance of the results.
- Mean and confidence interval: Means were reported with 95% confidence intervals to estimate the true mean of an entire population from the sample population provided for each study.

Conclusions

- Participants generally agreed that self-driving cars should sacrifice their passengers for the greater good, as indicated by high ratings for morality. However, they were unwilling to buy those cars for themselves. This social dilemma might prevent the adoption of self-driving cars (which would likely lead to safer roads) because people would rather buy cars that protect them from harm.
- Participants were even less willing to buy a self-driving car that sacrificed its passengers if they imagined that a family member was present in the car as well.

- Participants were less likely to think that self-driving cars should sacrifice their own passenger if there was only one pedestrian that could be saved. However, if there was more than one pedestrian, participants tended to think that it was moral for the self-driving car to sacrifice its passenger, with their approval ratings increasing with the number of pedestrians.
- Participants did not think that these types of moral decisions should be enforced by the government, as indicated by low approval ratings for both human and computer drivers. As further support for this conclusion, participants were significantly more willing to buy an unregulated autonomous vehicle than a regulated one.

LEARNING STANDARDS ALIGNMENT

The following tables provide an overview of the learning standards covered by this article, including the A Framework for K-12 Science Education (Framework), Common Core State Standards English Language Arts-Literacy (CCSS ELA), Common Core State Standards Statistics and Probability (CCSS HSS), AP Science Practices, and Vision and Change for Undergraduate Education. Where applicable, activities and information will be marked with specific standards to which they are linked.

A Framework for K-12 Science Education		
Science and Engineering Practices	Disciplinary Core Ideas	Crosscutting Concepts
Asking questions and defining problems (SEP1) Asking questions to determine relationships, including quantitative relationships, between independent and dependent variables. Define a design problem that involves the development of a process or a system with interacting components and criteria, with constraints that may include social, technical, and/or environmental considerations. **Analyzing and interpreting data (SEP4)** Analyze data using tools, technologies, and/or models (e.g. computational, mathematical) in order to make valid and reliable scientific claims or determine an optimal design solution. Compare and contrast various types of data sets (e.g., self-generated, archival) to examine consistency of measurements and observations.	**ETS2.B: Influence of engineering, technology, and science on society and the natural world** Not only do science and engineering affect society, but society's decisions (whether made through market forces or political processes) influence the work of scientists and engineers. These decisions sometimes establish goals and priorities for improving or replacing technologies; at other times they set limits, such as in regulating the extraction of raw materials or in setting allowable levels of pollution from mining, farming, and industry.	**Cause and effect** Empirical evidence is required to differentiate between cause and correlation and to make claims about specific causes and effects.

Science in the Classroom AAAS

Common Core State Standards English Language Arts-Literacy		
Key Ideas and Details	Craft and Structure	Integration of Knowledge and Ideas
RST.9-10.1 Cite specific textual evidence to support analysis of science and technical texts, attending to the precise details of explanations or descriptions. **RST.9-10.2** Determine the central ideas or conclusions of a text; trace the text's explanation or depiction of a complex process, phenomenon, or concept; provide an accurate summary of the text. **RST.11-12.1** Cite specific textual evidence to support analysis of science and technical texts, attending to important distinctions the author makes and to any gaps or inconsistencies in the account. **RST.11-12.2** Determine the central ideas or conclusions of a text; summarize complex concepts, processes, or information presented in a text by paraphrasing them in simpler but still accurate terms.	**RST.9-10.4** Determine the meaning of symbols, key terms, and other domain-specific words and phrases as they are used in a specific scientific or technical context relevant to grades 9-10 texts and topics. **RST.9-10.5** Analyze the structure of the relationships among concepts in a text, including relationships among key terms (e.g., force, friction, reaction force, energy). **RST.9-10.6** Analyze the author's purpose in providing an explanation, describing a procedure, or discussing an experiment in a text, defining the question the author seeks to address. **RST.11-12.4** Determine the meaning of symbols, key terms, and other domain-specific words and phrases as they are used in a specific scientific or technical context relevant to grades 11-12 texts and topics. **RST.11-12.5** Analyze how the text structures information or ideas into categories or hierarchies, demonstrating understanding of the information or ideas. **RST.11-12.6** Analyze the author's purpose in providing an explanation, describing a procedure, or discussing an experiment in a text, identifying important issues that remain unresolved.	**RST.9-10.8** Assess the extent to which the reasoning and evidence in a text support the author's claim or a recommendation for solving a scientific or technical problem. **RST.9-10.9** Compare and contrast findings presented in a text to those from other sources (including their own experiments), noting when the findings support or contradict previous explanations or accounts. **RST.11-12.8** Evaluate the hypotheses, data, analyses, and conclusions in a science or technical text, verifying the data when possible and corroborating or challenging conclusions with other sources of information. **RST.11-12.9** Synthesize information from a range of sources (e.g., texts, experiments, simulations) into a coherent understanding of a process, phenomenon, or concept, resolving conflicting information when possible.

Science in the Classroom AAAS

AP Science Standards	
AP Science Practices	AP Statistics Content Standards
Science Practice 2 (SP2) The student can use mathematics appropriately. **Science Practice 5 (SP5)** The student can perform data analysis and evaluation of evidence. **Science Practice 7 (SP7)** The student is able to connect and relate knowledge across various scales, concepts, and representations in and across domains.	**Data Collection** Data must be collected according to a well-developed plan if valid information on a conjecture is to be obtained. This plan includes clarifying the question and deciding upon a method of data collection and analysis.

Connections to the Nature of Science	
Vision and Change for Undergraduate Biology Education Core Competencies and Disciplinary Practices	A Framework for K-12 Science Education Understandings About the Nature of Science
Ability to use quantitative reasoning Understand how to apply quantitative approaches (statistics, quantitative analysis of dynamic systems, and mathematical modeling) to solve a variety of problems. **Ability to understand the relationship between science and society** Identify social and historical dimensions of scientific practice: evaluating the relevance of social contexts to scientific problems, developing scientific applications to solve societal problems, evaluating ethical implications of scientific research.	**Scientific addresses questions about the natural and material world** Not all questions can be answered by science. Science and technology may raise ethical issues for which science, by itself, does not provide answers and solutions. **Scientific knowledge is based on empirical evidence** Scientific arguments are strengthened by multiple lines of evidence supporting a single explanation. **Scientific investigations use a variety of methods** Scientific investigations use a variety of methods, tools, and techniques to revise and produce new knowledge