Social interactions between human and machine drivers

Peng Liu (刘鹏)
Zhejiang University
pengliu86@zju.edu.cn

Education

QS rank: 334		QS rank: 17
Bachelor	Visiting Student	PhD in Human Factors
Industrial Engineering	Human Factors and Applied Cognition	Industrial Engineering
2005.09–2009.06	2013.08–2013.12	2009.09–2014.07

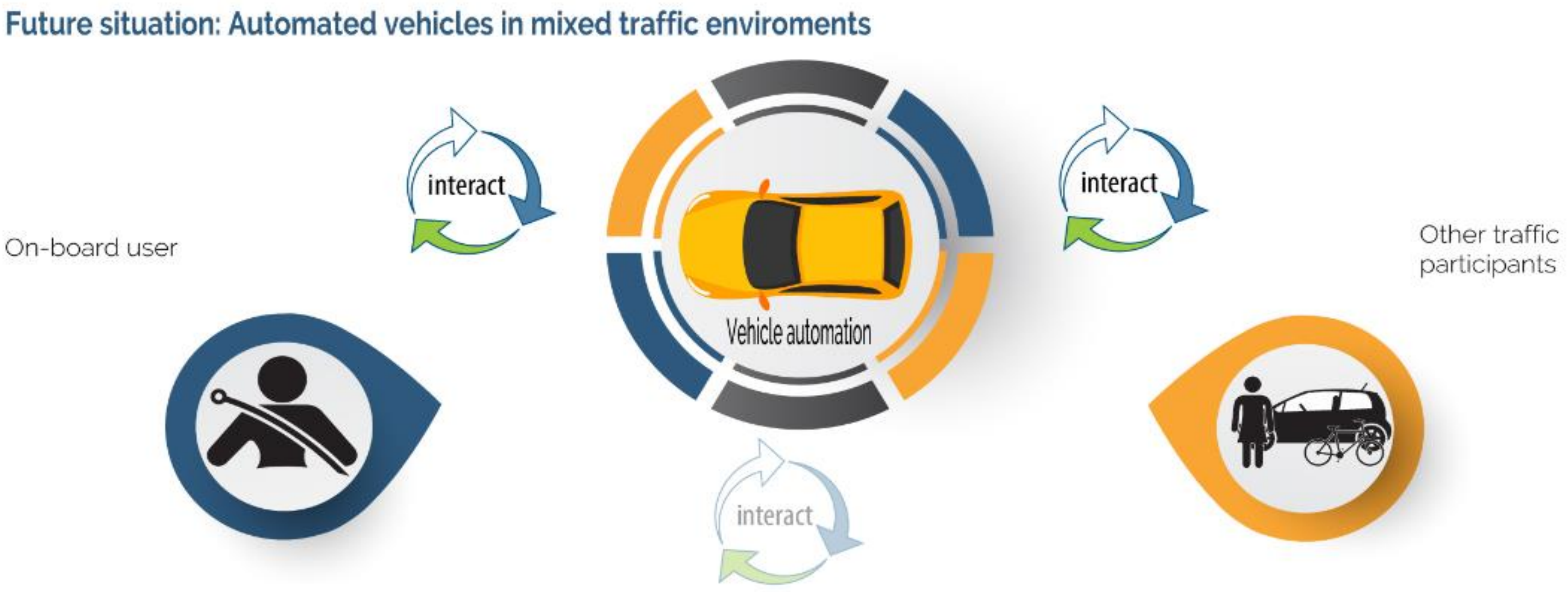
Working experience

QS rank: 334		QS rank: 45
Assistant/Associate Prof.	Visiting Researcher	Researcher
Industrial Engineering	Experimental Psychology	Psychological Sciences
2015.11–2021.03	2017.11–2018.11	2021.04–present

Source: interAct (https://www.interact-roadautomation.eu/)

Roads are “inherently social”

Driving is a social activity



Source: interAct (https://www.interact-roadautomation.eu/)

Reliable, efficient, and smooth social interaction between human and machine drivers (AVs) in mixed traffic is the key.

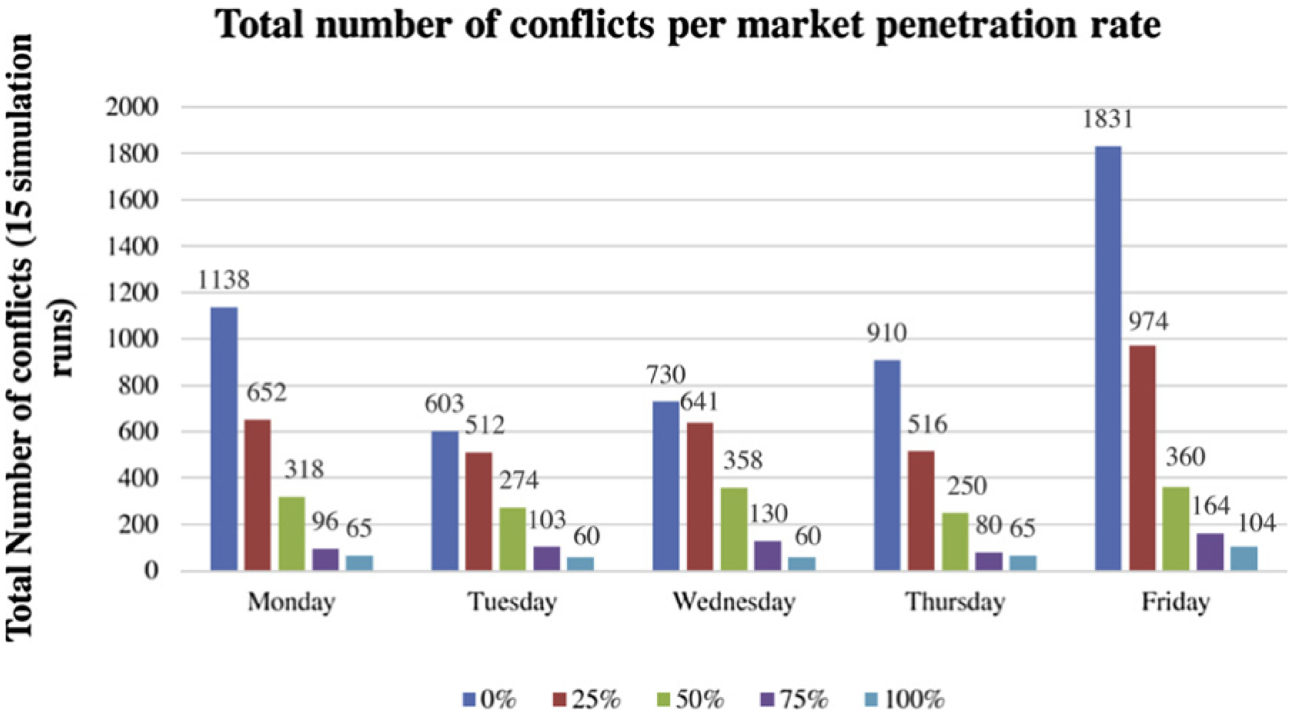
So, shall we worry about human-AV social interactions?

Machine drivers are imagined or supposed to be a super driver devoid of human weaknesses such as attention lapse, distraction, fatigue, drunk and alcohol-impaired driving, etc.

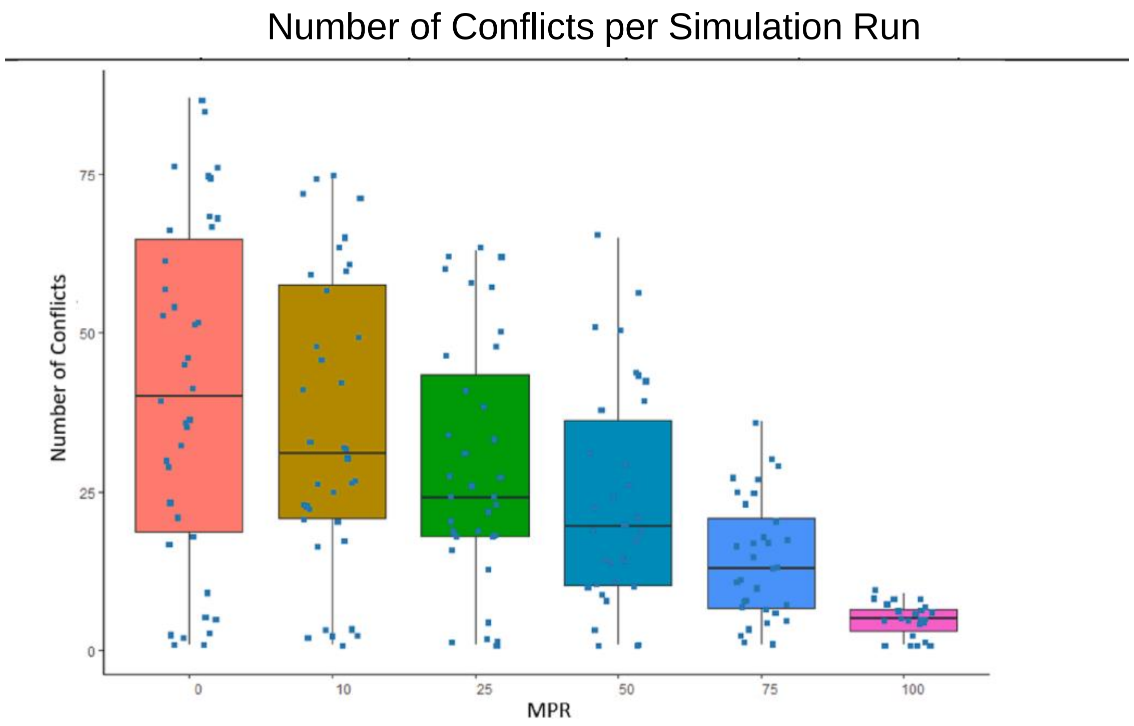
They have many advantages: shorter reaction times, shorter following headways, more predictable, etc.

They are promised to eliminate 90% of accidents.

Assume AVs as a super driver and with so many capabilities, dozens of simulation-based assessments reported the safety and operational benefits of AVs (with connectivity)



Papadoulis et al., 2019, AAP

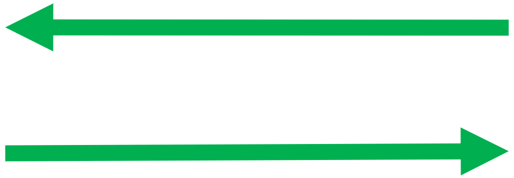


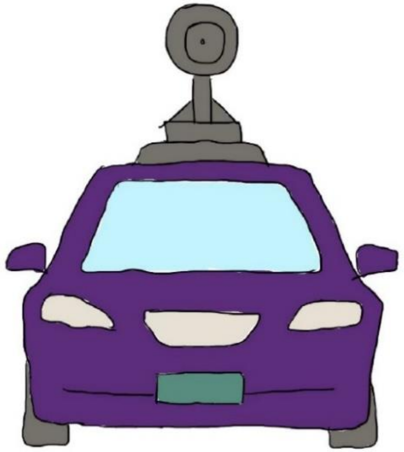
Mousavi et al., 2021, AAP

Given these huge benefits, why shall we worry about human-AV social interactions?

- Collectively “dance” with other drivers following “different implicit and sometime explicit rules”
- Adaptable, exhibit satisficing rather than optimizing driving behaviors
- Aggressive, moderate, and conservative (defensive)

- As strangers, machines do not know “implicit rules” and are hard to learn “driving etiquette”
- Strict rule-followers
- Conservative (defensive), overly cautious





mix like oil and water

Refs: Hancock et al. (2019); Nyholm & Smids (2020)

Would you bully a driverless car or show it respect? (Wall; BBC, 2016)

Humans will bully mild-mannered autonomous cars (Condliffe; MIT Technology Review, 2016)

A slashed tire, a pointed gun, bullies on the road: Why do Waymo self-driving vans get so much hate? (Randazzo; Arizona Central, 2018)

Uber says people are bullying its self-driving cars with rude gestures and road rage (Hamilton; Business Insider, 2019)

AV bullying could be a random event, because:

- There will always be some people who drive aggressively to get an edge. So, these people would also “bully” AVs.

But what if AV bullying is not random?

Human-AV social interactions

Etiquette equality hypothesis (human-machine inter. = human-human inter.)

- "Media Equation" paradigm (Reeves & Nass, 1996)
- "Computers-Are-Social-Actors" paradigm (CASA; Nass & Moon, 2000)
- Social scripts to human-human interaction are applied *mindlessly* as a heuristic shortcut in human-computer interaction.

Etiquette inequality hypothesis (human-machine inter. ≠ human-human inter.)

- "Media Inequality" paradigm (Shechtman & Horowitz, 2003; Mou & Xu, 2017)
- Socially interacting with machines is not *mindless*, but *mindful*.
- People would *mindfully* realize that they are not real humans and so they have different cognitive, affective, and social responses to them.

Source: Li & Liu (in press, Book chapter)

Etiquette equality hypothesis

or

Etiquette inequality hypothesis

Our works

- Study 1: **Liu***, Du, Wang*, Ju. (2020). Ready to bully automated vehicles on public roads? Accident Analysis & Prevention, 137, 105457.
- Study 2: **Liu***, Zhai, Li. (accepted). Is it OK to bully automated cars? Accident Analysis & Prevention.

} Driving aggression (bullying)

- Study 3: Li, Wang, Liu, Yuan, **Liu*** (2022). Sharing the roads: Robot drivers (vs. human drivers) might provoke greater driving anger when they perform identical annoying driving behaviors. International Journal of Human-Computer Interaction. 38 (4): 309–323.

Driving anger

- Study 4: Li, **Liu*** (in press). Etiquette equality or inequality? Drivers' intention to be polite to automated vehicles in mixed traffic. In Duffy, V. G., Landry, S. L., Lee, J. D., & Stanton, N. A. (eds.), Human-Automation Interaction: Transportation. Springer.

Driving politeness

An imaginary story

Imagine two cars waiting at a junction: an automated vehicle (AV) and a human-driven vehicle.

The AV has the right of way, but the human driver proceeds anyway.

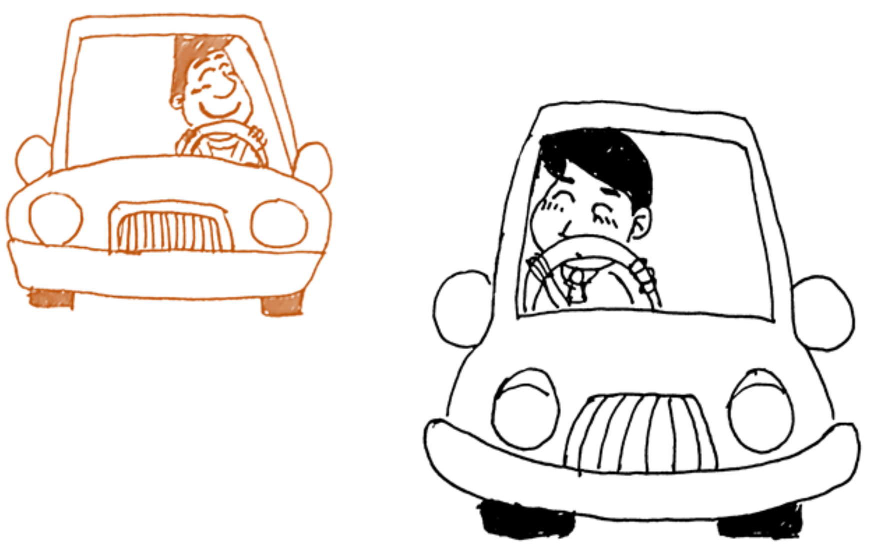
Then, the AV has to stop itself and let the aggressive human driver go.

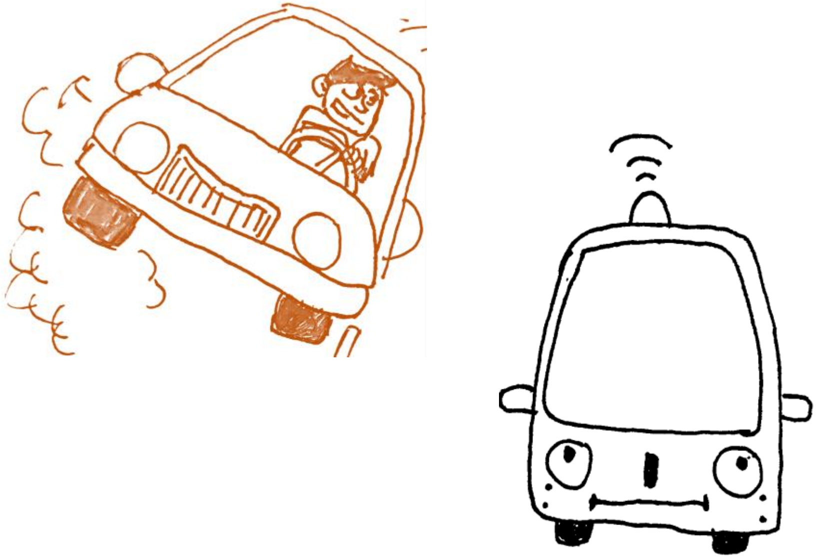
The human driver says, “I will be overtaking all the time because the automated car will be sticking to the rules.”

Source: Liu et al., 2020, AAP

Study 1: Ready to bully automated vehicles on public roads?

- To examine whether human drivers are more likely to “bully” AVs than to bully other human drivers.
- To explore why.





Methodology

- Bullying behaviors refer to any kind of improper or violent behaviors intended to cause damage or impede the operation of other vehicles.
- Bullying Intention Questionnaire (BIQ; 11 behaviors) based on the Driver Behavior Questionnaire (DBQ; Reason et al., 1990)
- Procedure:
  - For each bullying behavior, participants read a traffic scenario and then indicated their likelihood to perform the bullying behavior: “Imagine the following scenario. On an urban road, there are three lanes in the same direction. You are driving a car in the middle lane. Another car is driving slowly in front of you. If you stay in the lane, you will have to slow down.” (Note: passing on right is not friendly).
  - They answered: “In this situation, based on your driving habit, how likely are you to overtake this car on the right side?” on a ten-point scale (1 = *very low*; 10 = *very high*)

Methodology

- Attitude toward AVs on a five-point scale ("*very negative*," "*negative*," "*neutral*," "*positive*," and "*very positive*")
- Risk-benefit perception of AVs to society on a five-point scale ("*risks far outweigh benefits*," "*risks outweigh benefits*," "*risks and benefits are the same*," "*benefits outweigh risks*," and "*benefits far outweigh risks*")
- 2 (Scenario: Human driver vs. AV) * 2 (Nationality: China and South Korea) between-subjects design; 495 participants from China and 503 participants from South Korea

Results

- Greater intention to bully AVs than to bully other human drivers

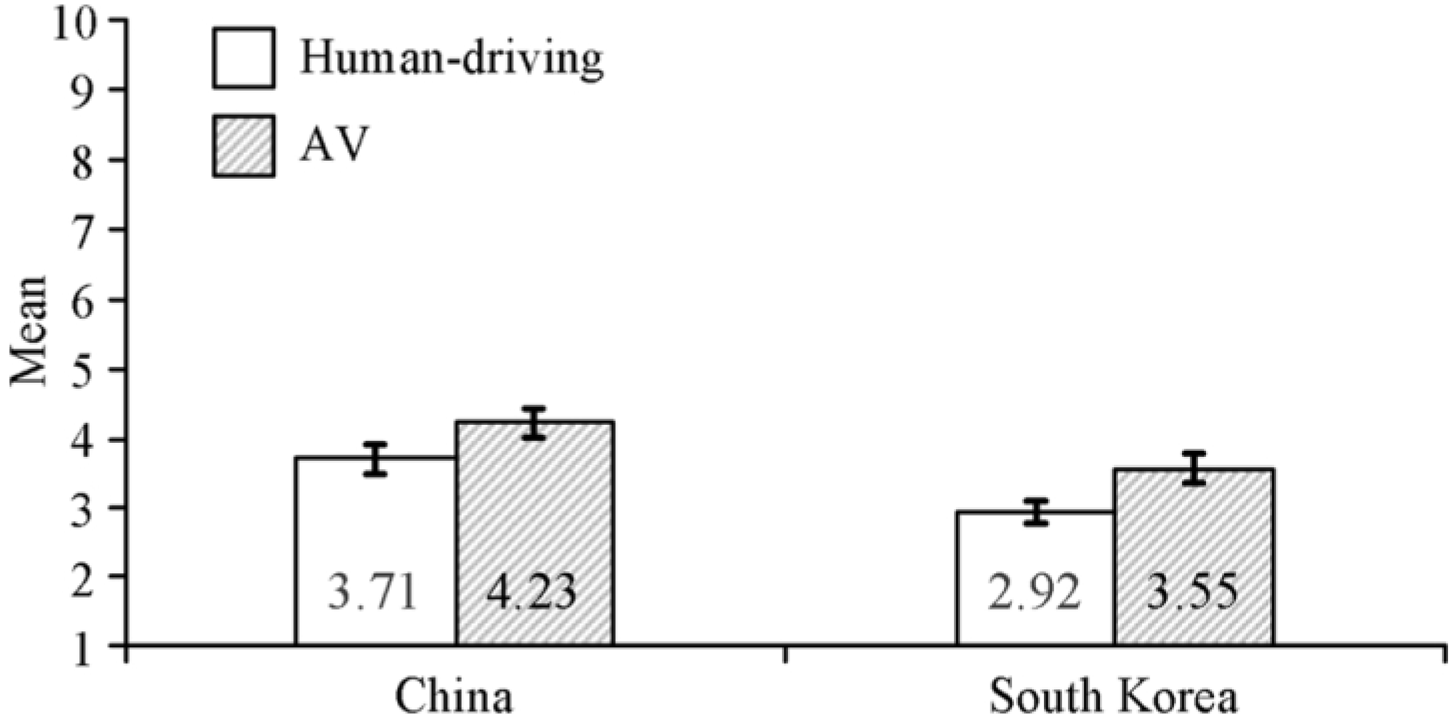


Fig. 1. Mean values of the overall bullying intention in the two scenarios (human-driving and AV) and countries (China and South Korea). Error bars = ± 2 standard error.

Results

- Female (vs. male) reported a lower intention to perform certain bullying behaviors

Results of ANCOVA in the Chinese survey.

Outcome variable	Scenario			Gender		
	B	p	η_p^2	B	p	η_p^2
BI1	0.52	.028	.010	−0.28	.232	.003
BI2	0.74	< .001	.023	−0.64	.003	.017
BI3	0.69	.001	.021	−0.54	.011	.013
BI4	0.47	.026	.010	−0.60	.004	.017
BI5	0.70	.001	.021	−0.81	< .001	.029
BI6	0.40	.045	.008	−0.76	< .001	.029
BI7	0.61	.008	.014	−0.41	.075	.006
BI8	0.44	.036	.009	−0.34	.104	.005
BI9	0.67	.002	.019	−0.54	.014	.012
BI10	0.98	< .001	.044	−0.71	< .001	.023
BI11	0.04	.850	.000	−0.73	< .001	.022
Bullying intention	0.57	< .001	.027	−0.58	< .001	.028

Results

- Pooled data: intention to bully AVs correlated with attitude toward AVs (r = −.087, p = .047) but not with risk-benefit perception (r = −.032, p = .468).
- China: intention to bully AVs correlated with attitude toward AVs (r = −.165, p = .007) and risk-benefit perception (r = −.122, p = .047)
- South Korea: none of them were significant (ps > .05).

- People's attitude and perceptions of AVs might not be robust predictors to their intention to drive aggressively toward AVs. That is, general attitude toward an object might weakly influence specific behaviors toward this object: **weak (general) attitude–(specific) behavior correspondence**.

Why could people be more likely to “bully” AVs?

General attitude and perception toward AVs might be not a good account.

- weak (general) attitude–(specific) behavior correspondence

We need to focus on bullying behaviors *per se*.

Why could people be more likely to “bully” AVs?

Cognitive

- AVs are conservative and risk-averse; so, bullying AVs is less risky
- ……

Affective

- Bullying AVs make some anticipates less negative affect
- ……

Moral

- Bullying machine drivers is less immoral
- ……

Study 2: Is it OK to bully automated cars?

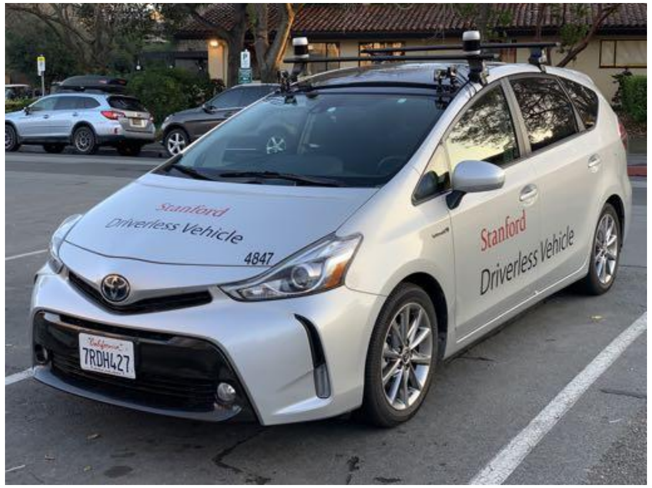
- To understand the observers' appraisals of aggression toward AVs versus human drivers.
- To indirectly explain why people would be more likely to drive aggressively towards AVs.

Identical aggression toward AVs versus human divers is less negatively appraised

↓

Greater intention to drive aggressively toward AVs

Moore et al. (2020): griefing behaviors of a Wizard-of-Oz driverless car on roads



A car ahead suddenly brakes fully and continuously

Modified from Moore et al. (2020)

Hypotheses

- H1: More accept aggression toward AV (vs. human driver)
- H2: Perceive fewer risks of aggression toward AV
- H3: Report less negative affect evoked from aggression toward AV
- H4: Judge aggression toward AV as less immoral
- H5: Differences in appraisals of the aggression toward the AV and human driver assumed in H1−H4 are more significant when the bullied victim identity salience is high than when the victim identity salience is low.

Methodology

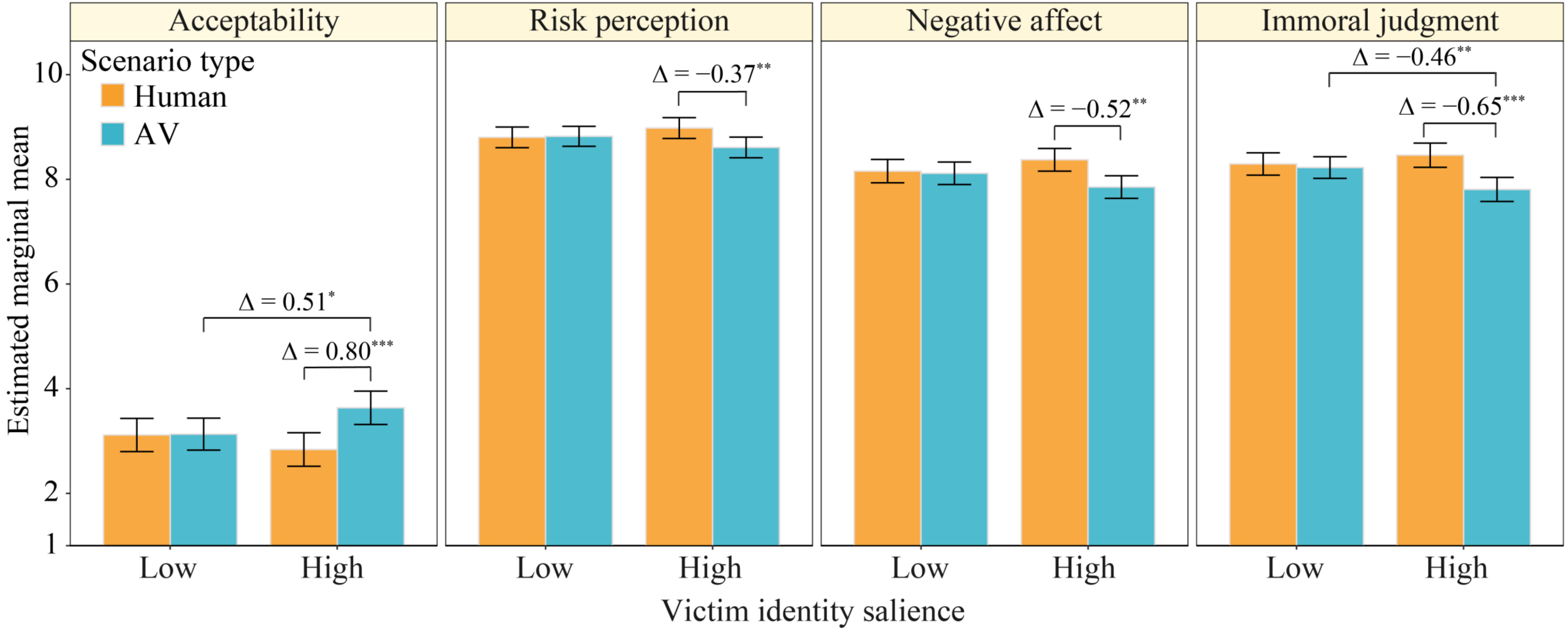
- 2 (scenario: human driver vs. AV) × 2 (victim identity salience: low vs. high) between-subjects design (N = 956)
- We informed participants that this video is from a human driver's car or an automated driving car
- Our video cannot visually make the bullied human driver's car and AV different. Thus, we manipulated the salience of victim identity in the measure design:
  - Low salience: participants evaluated the aggressive behavior
  - High salience: participants evaluated the aggressive behavior toward the AV or the human driver

Methodology

- Evaluated on a 10-point scale (1 = *very low*; 10 = *very high*)
- Acceptability and risk perception measured by a single-item design
- Negative affect and immoral judgment measured by a multi-item design
- Items were adapted from previous studies

Factor	Item	Factor loading		ITC
		I	II	
Negative affect	Anger	.06	**.81**	.73
	Disgust	.06	**.81**	.73
Immoral judgment	Morally prohibited	**.79**	.06	.78
	Morally wrong	**.84**	.02	.79
	Morally blameworthy	**.72**	.19	.80

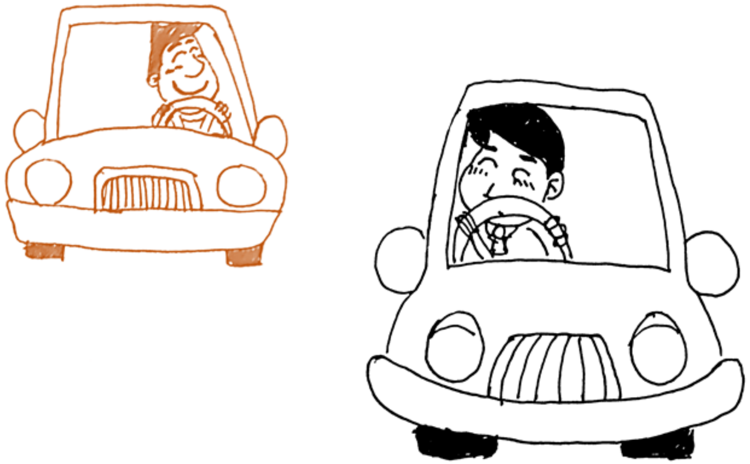
Result

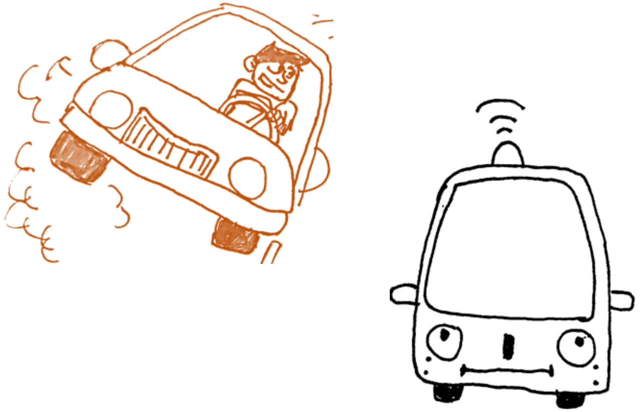


When the victim identity of this behavior was salient during participants' appraisals, they appraised this behavior toward AVs less negatively

Lance Eliot "An Inconvenient Truth: Human Drivers and Autonomous Cars Mix Like Oil And Water" (Forbes, 2019):

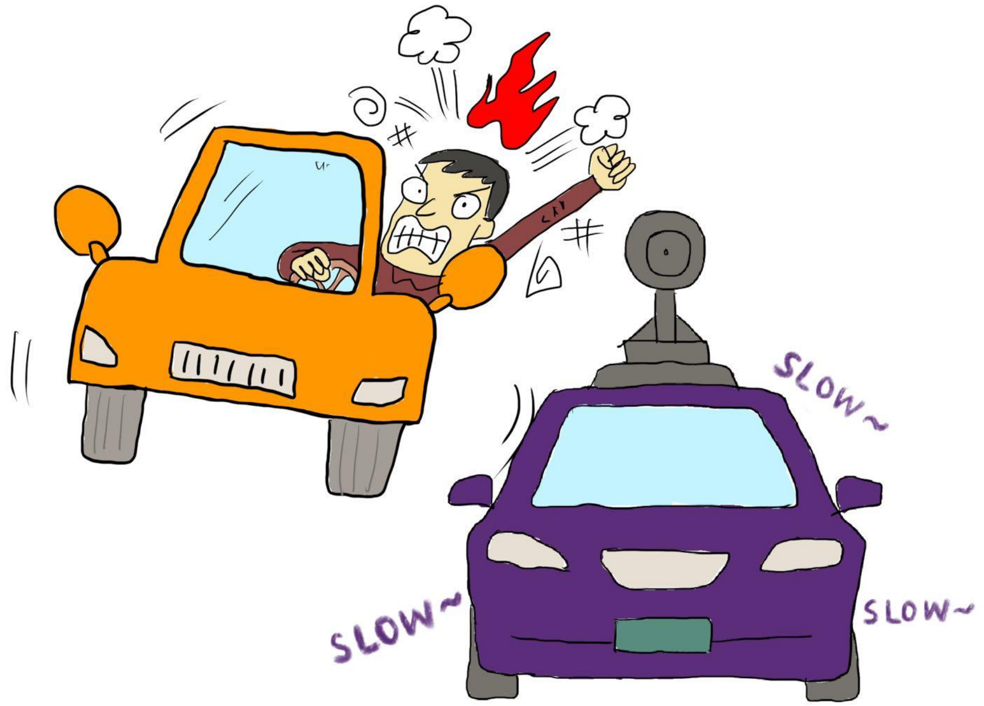
"For those weathered human drivers that have endured being around today's autonomous cars, they have also discovered that they can outrun and outfox the driverless car. Why wait behind the slowpoke self-driving car when you can skip around it, and it won't seemingly care. Why not proceed to cut-off the driverless car so that you can make a turn sooner, and therefore not need to get stuck waiting for the stilted AI to get the self-driving car to agonizingly slowly make that same turn."





If a car is driving slowly in front of you, what will you do?

What if the slowing car is an AV?



Current AVs could irritate human drivers

More than 60% of AV crashes were rear-end ones (Petrović et al., 2020), compared as 14% of conventional cars were rear-ended by another car (Teoh & Kidd, 2017)

These over-represented rear-end collisions were probably because AVs behaved in ways that human drivers did not anticipate or wish them to behave

Other major AV crashes were side-swipes (Petrović et al., 2020), many of which appeared to "involve human drivers frustrated at getting stuck behind a slow or stopped AV, trying to overtake it, and not quite making it" (Stewart, 2018)

Study 3: Sharing the roads: Robot drivers (vs. human drivers) might provoke greater driving anger when they perform identical annoying driving behaviors

- To explore whether AVs' aberrant behaviors provoke higher driving anger than do identical behaviors performed by human drivers
- To explore why

Methodology

- Driving anger scale (DAS; adapted from Deffenbacher et al., 1994) with a scale (*not at all* = 1, *a little* = 2; *some* = 3; *much* = 4; *very much* = 5)

Item	Description
DA1	A car is weaving in and out of traffic
DA2	A car driving slowly on a mountain road will not pull over and let people by
DA3	A car backs right out in front of you without looking
DA4	A car runs a red light
DA5	When you are ready to overtake a car, it speeds up and thus impedes your overtaking
DA6	A car is slow in parking and holding up traffic
DA7	A car honks at you about your driving
DA8	A car changes its lane and overtakes your car when it drives really close to your car

DA1– DA7 were adapted from Deffenbacher et al.'s (1994) short-form DAS, while DA8 was adapted from Liu et al. (2020a).

Methodology

- Mind perception (perception of whether AVs have human-like minds) with two dimensions: agency and experience (Gray et al., 2007)

		Factor loading				
Factor	Item	I	II	ITC	*M*	*SD*
Experience (α = 0.90)	Feeling pain	0.82	0.06	0.77	1.86	1.13
	Feeling fear	0.81	0.14	0.78	2.16	1.31
	Feeling angry	0.91	0.08	0.84	2.00	1.17
	Feeling pleasure	0.76	0.22	0.73	2.10	1.25
Agency (α = 0.77)	Self-control	0.12	0.68	0.59	3.30	1.21
	Planning	−0.07	0.75	0.57	3.53	1.16
	Acting morally	0.20	0.71	0.63	3.09	1.31
	Thinking	0.39	0.50	0.48	2.61	1.30

Results

- Participants (N = 622) in the AV scenario reported higher anger than those in the human-driving scenario
- Older participants reported experiencing lower driving anger
- No gender difference

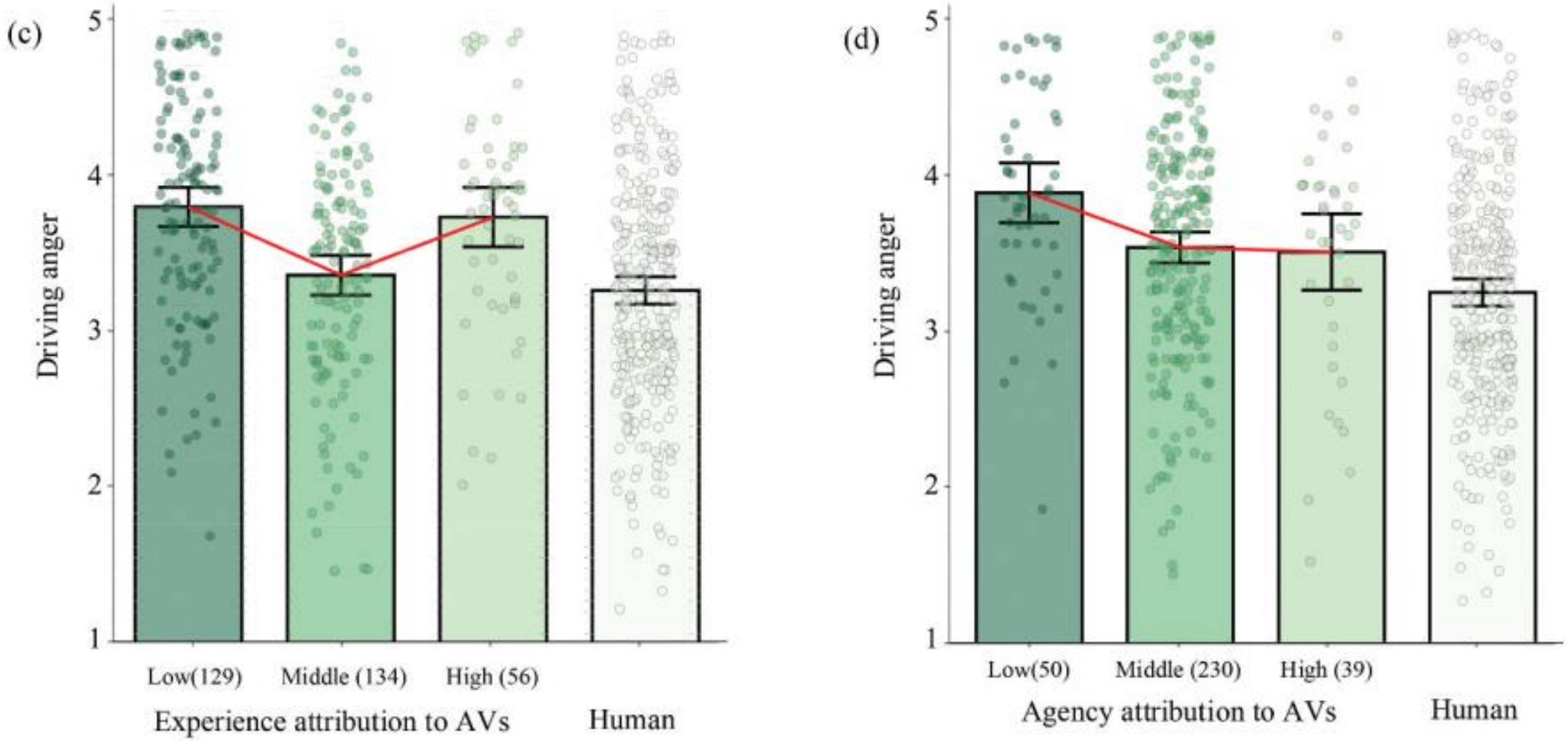
Table 4. Results of ANCOVA.

Outcome Variable	Scenario			Gender			Age		
	B	F	η_p^2	B	F	η_p^2	B	F	η_p^2
DA1	0.21*	5.27	.008	0.01	0.02	.000	−0.41**	10.57	.017
DA2	0.22**	6.82	.011	−0.12	2.09	.003	−0.20	2.80	.005
DA3	0.25**	8.38	.013	0.04	0.20	.000	−0.39**	10.88	.017
DA4	0.56***	31.84	.049	0.02	0.06	.000	0.09	0.45	.001
DA5	0.42***	20.84	.033	−0.19*	4.07	.007	−0.56***	19.02	.030
DA6	0.30**	10.58	.017	−0.17	3.31	.005	−0.56***	19.53	.031
DA7	0.40***	19.77	.031	−0.12	1.59	.003	−0.78***	37.34	.057
DA8	0.30**	9.41	.015	0.18	3.55	.006	−0.48***	12.84	.020
Total driving anger	0.33***	28.59	.044	−0.04	0.45	.001	−0.41***	22.50	.035

B = unstandardized coefficients; Scenario: human = 0, AV = 1; Gender: male = 0, female = 1; Age: ≤ 30 = 0, > 30 = 1. *p < .05; **p < .01; ***p < .001.

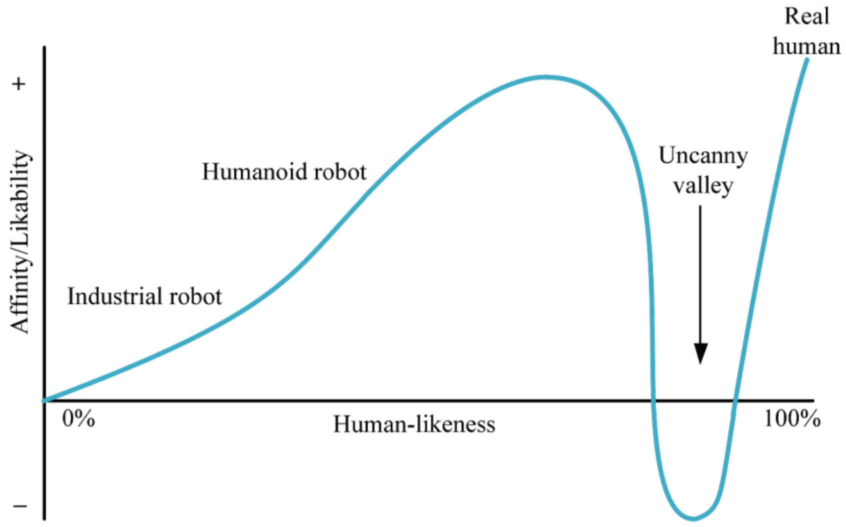
Results

- Negative relationship between agency attribution and driving anger (d)
- U-relationship between experience attribution and driving anger (c)

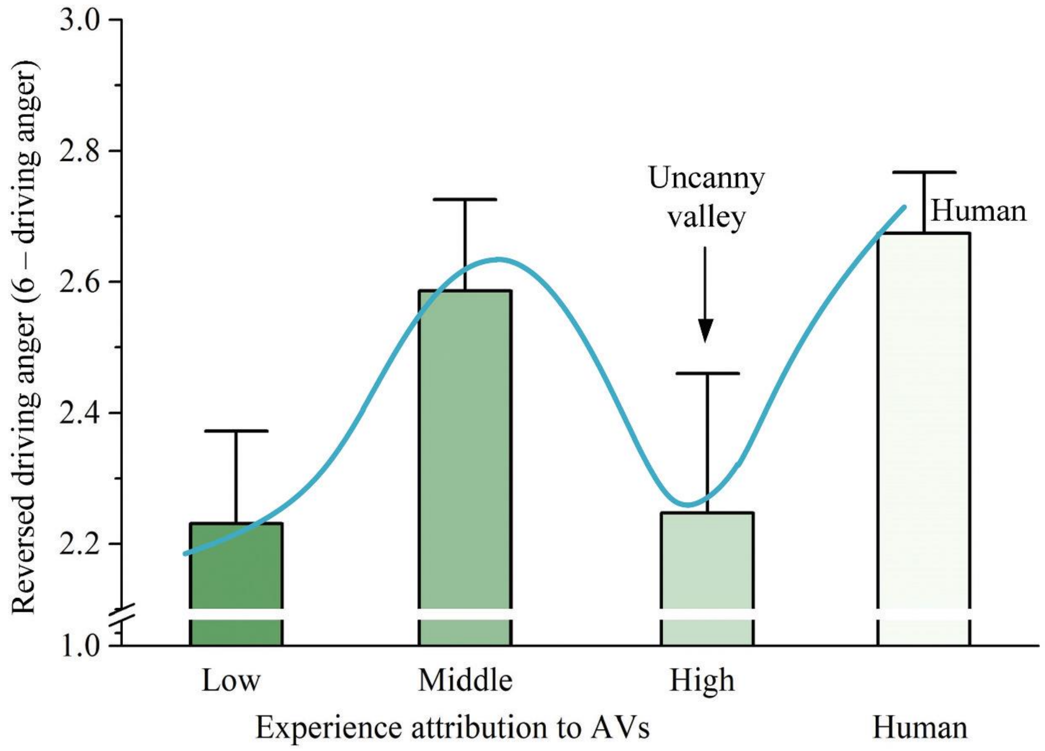


Results

- U-inversed relationship between experience attribution to AVs and reversed driving anger (6 – driver anger) may reflect the “uncanny valley”

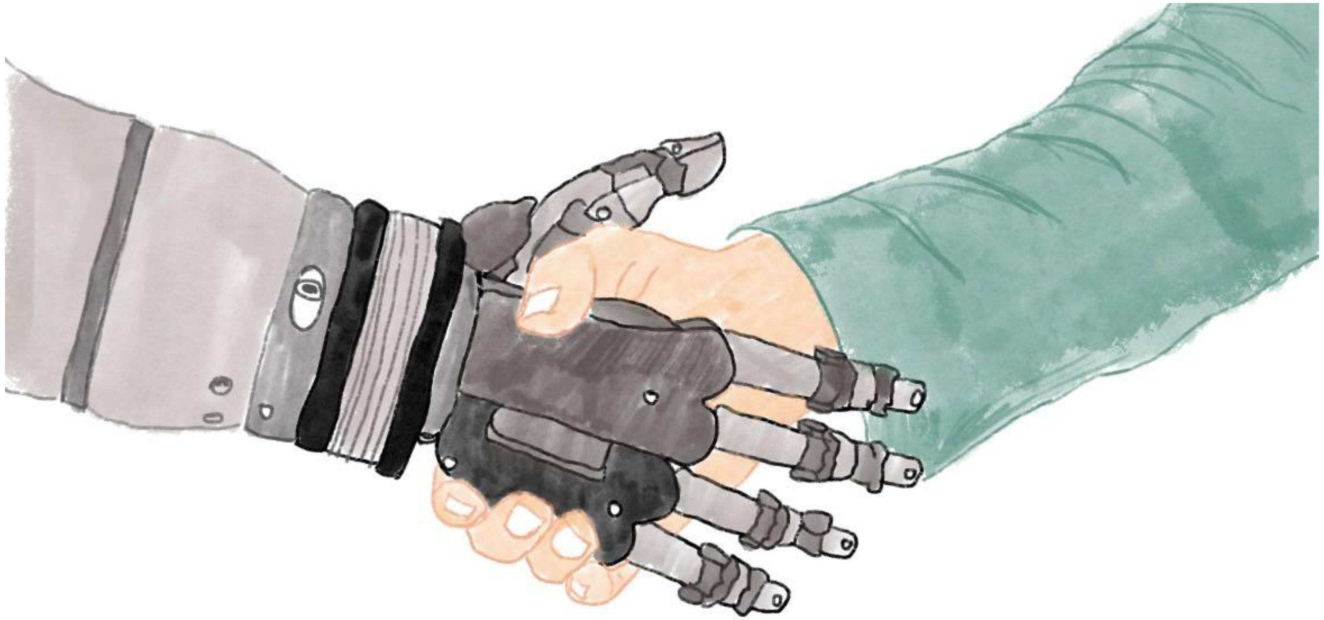


“Uncanny valley”: increasingly human-like appearance of robots would lead to increases in the degree to which people like them only up to a certain point, after which robots appear too human and become unnerving



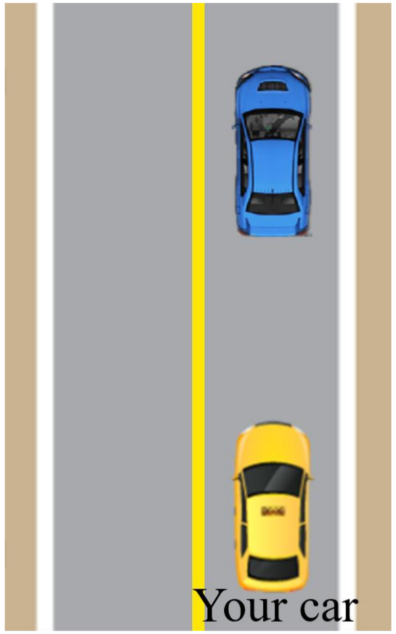
Study 4: Etiquette equality or inequality? Drivers' intention to be polite to automated vehicles in mixed traffic

- To explore human drivers' intention to negotiate with AVs versus other human drivers.



Methodology

- Our polite driving behavior questionnaire (PDBQ) was based on the positive driver behaviors scale (Özkan & Lajunen, 2005) and multidimensional driving style inventory (Taubman-Ben-Ari et al., 2004); with nine traffic scenarios, e.g.:



There is a car ahead of you in the lane.

E1: how likely are you to avoid closely following and not to disturb the car ahead?

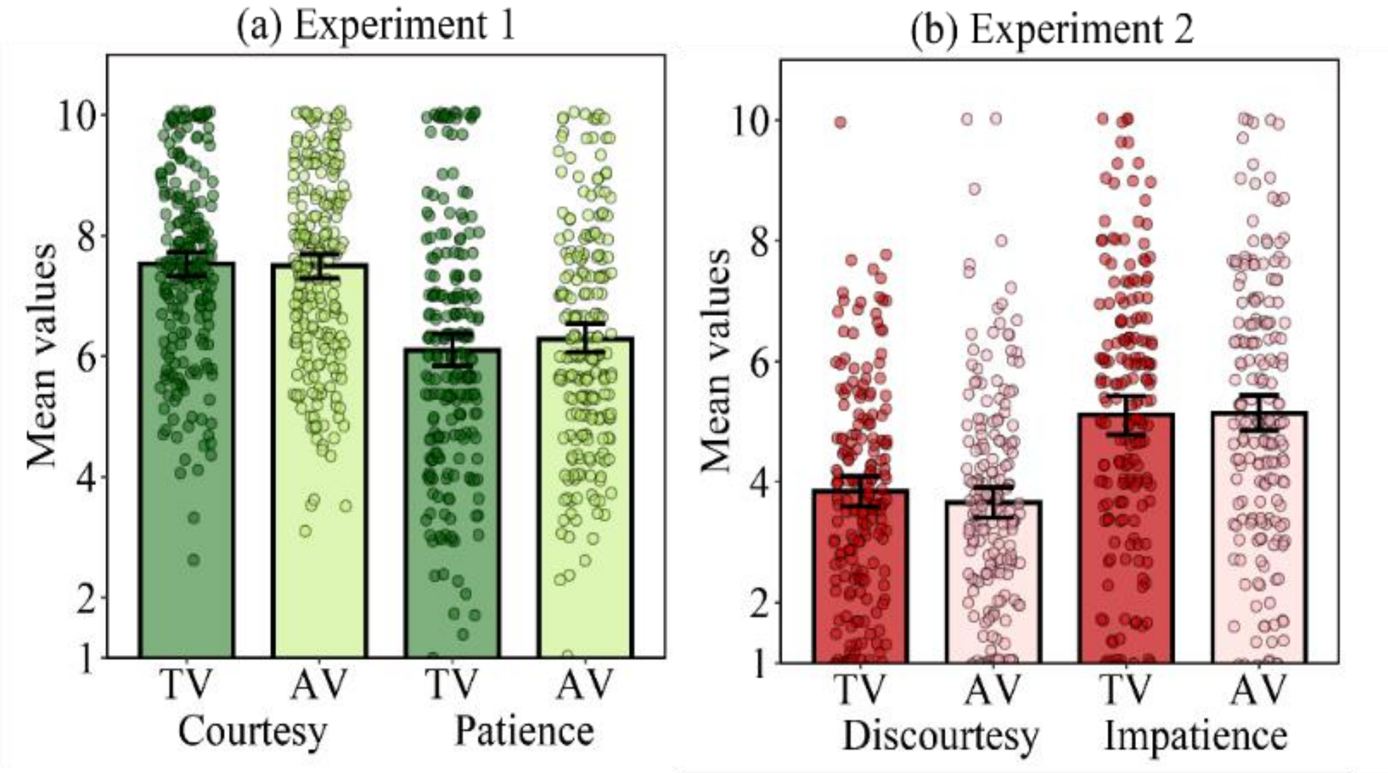
E2: how likely are you to keep closely following the car ahead?

→ To measure driving politeness in Experiment 1

→ To measure driving impoliteness in Experiment 2

Results

- Nine traffic scenarios were factored into two constructs: courtesy and patience in Experiment 1 (or discourtesy and impatience in Experiment 2)



TV: traditional vehicle; AV: automated vehicle

Results

- Participants' intentions to perform polite behaviors (courtesy and patience) or impolite behaviors (discourtesy and impatience) were similar when interacting with the two different cars, contrary to our Studies 1-3.
- Potential explanations:
  - In line with the etiquette equality hypothesis, participants regarded an AV as a normal car and applied the same driving etiquette and interactive behaviors to them when AVs perform common driving behaviors.
  - Or, the vehicle type (automated vs traditional car) was not a salient cue for influencing participants' intention to negotiate with other cars in our survey.
  - Or, the human-machine asymmetry exists only in cases where AVs are involved with negative driving behaviors.

Summary from our current works

- Human-machine asymmetry: AVs are treated more harshly
  - Human drivers' greater intention to drive aggressively toward AVs (Liu et al., 2020, AAP)
  - This difference could be associated with different appraisals of aggression toward AVs and human drivers. Aggression toward AVs was less negatively evaluated (Liu et al., accepted, AAP)
  - AVs' annoying driving behaviors might elicit greater driving anger (Li et al., 2022, IJHCI)
- However, participants expressed similar intentions to demonstrate polite driving behaviors while negotiating with AVs versus human drivers (Li & Liu, in press, book chapter).

Studies supporting our findings

- Lee et al. (2021, Technology in Society): vignette-based survey
  - Modify the Propensity for Angry Driving Scale (PADS; > 10 aversive traffic scenarios) and design Human-Driver PADS and Self-Driving Car PADS.

You are in the left-hand lane behind a self-driving car. When the left turn light is given, the car does not move. You tap on your horn to get the car's attention and it signals an obscene gesture towards you.

A: Which of the following thoughts about the car is most likely to go through your mind?

1) It intentionally did this to mess up the traffic.
2) It did not mean anything wrong, just a mistake.

B: In the above situation, please rate how much would you feel each of the following emotions on a scale of 1-5 where 1 = not at all, and 5 = very much.

	Not at all				Very much
Anger	1	2	3	4	5
Irritated or annoyed	1	2	3	4	5

C: How would you respond?

intended responses

1) Tap on your horn again.
2) Fume inside a bit, but do nothing.
3) Lay on your horn.
4) Lay on the horn and return an obscene gesture.

"participants judged the actions of self-driving cars as unintentional but chose to react more aggressively toward them in return"

"This finding is consistent with prior research that, in mixed traffic, human drivers exhibit more aggressive behaviors toward AVs than toward other fellow drivers [8,9]"

Human-machine asymmetry (Liu et al., 2020, AAP)

Studies supporting our findings

- Ma and Zhang (in press, Human Factors): video-based survey

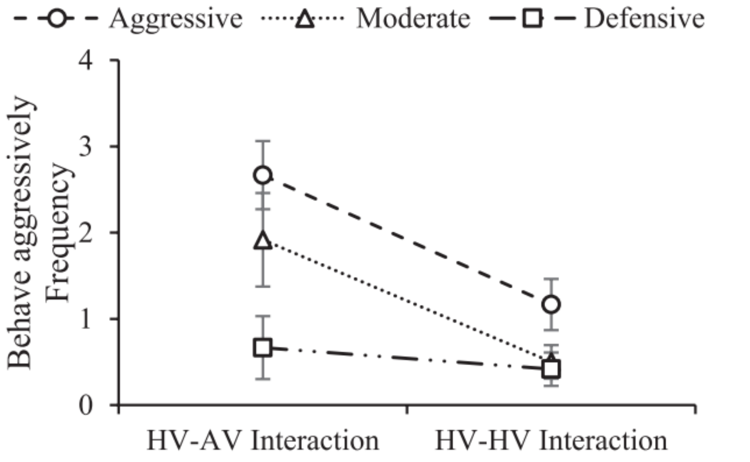
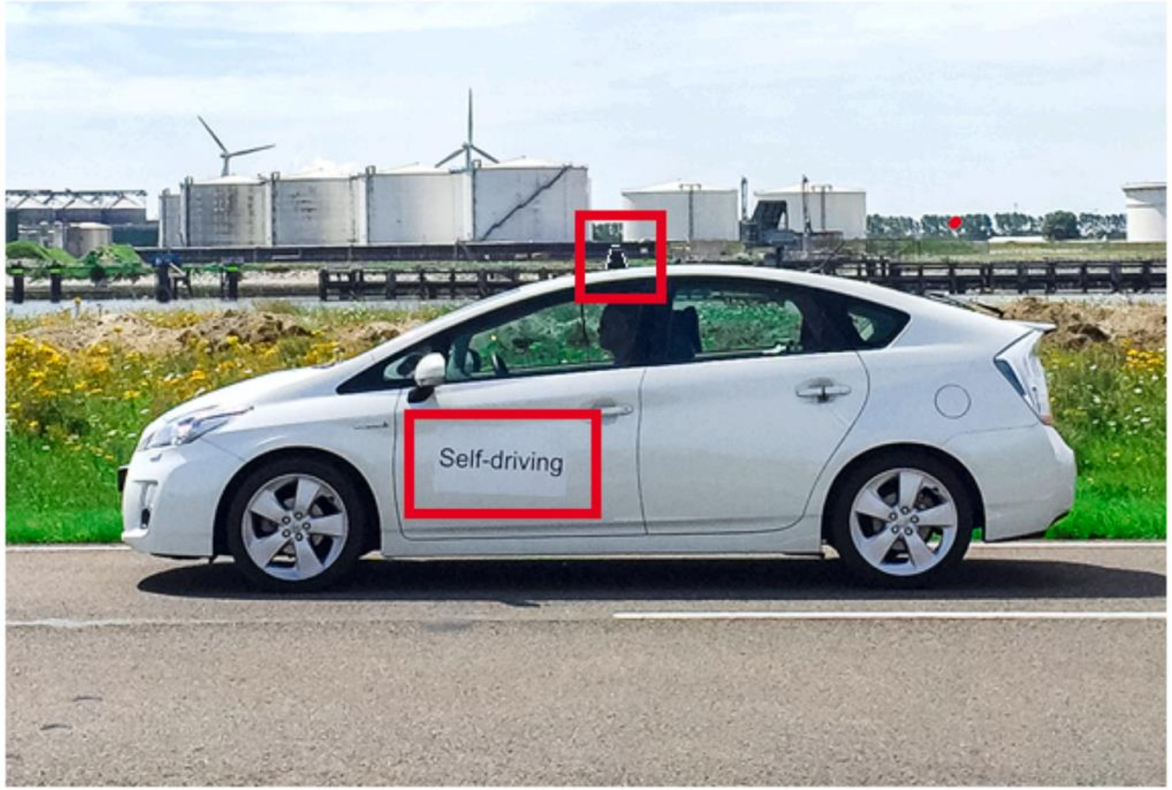


Figure 4. Effects of driving style and type of interaction on drivers' intention to behave aggressively (error bars: ± 1 SE).

"This finding expands on those of Liu et al. (2020). In their survey study, Liu et al. found drivers were more likely to behave aggressively in HV-AV interaction than in HV-HV interaction; however, they did not consider drivers' driving styles. **The findings of this study indicate that only aggressive and moderate drivers showed changes in their intentions to behave more aggressively in HV-AV interaction**"

Studies supporting our findings

- Soni, Reddy, et al. (2022, TR-F): “Wizard of Oz” AV + field experiment

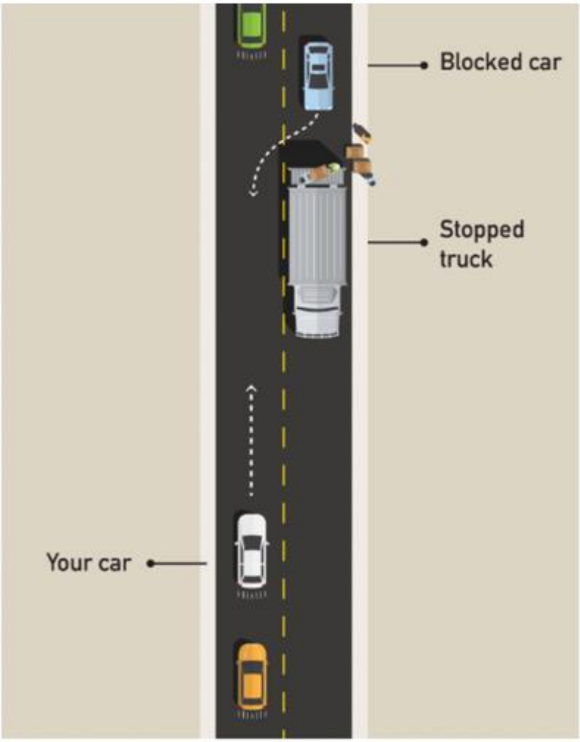


“The results show that human drivers adopted significantly smaller critical gaps when interacting with the approaching AV as compared to when interacting with the approaching HDV. Drivers also maintained a significantly shorter headway after overtaking the AV in comparison to overtaking the HDV.”

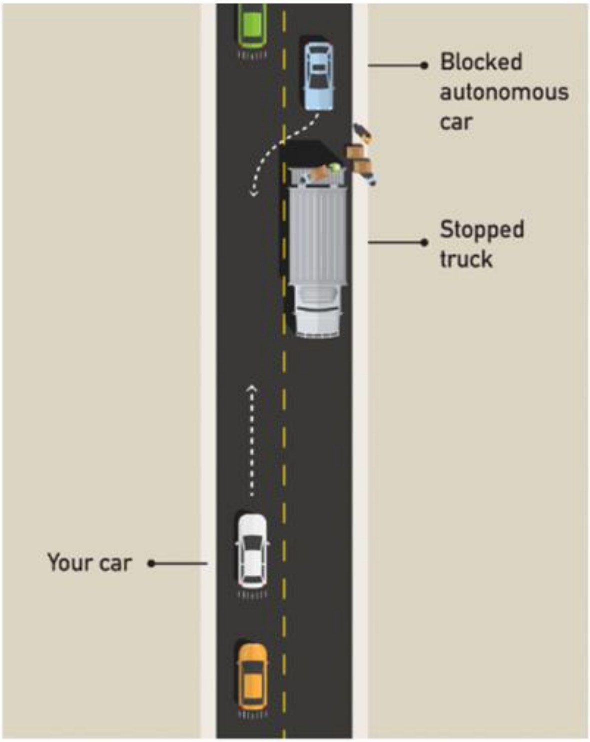
Humans’ behavioral adaptation and potential exploitation of AV technology

Studies without supporting our findings

- Wong (2019, AutoUI'19): negotiate with automated vehicle



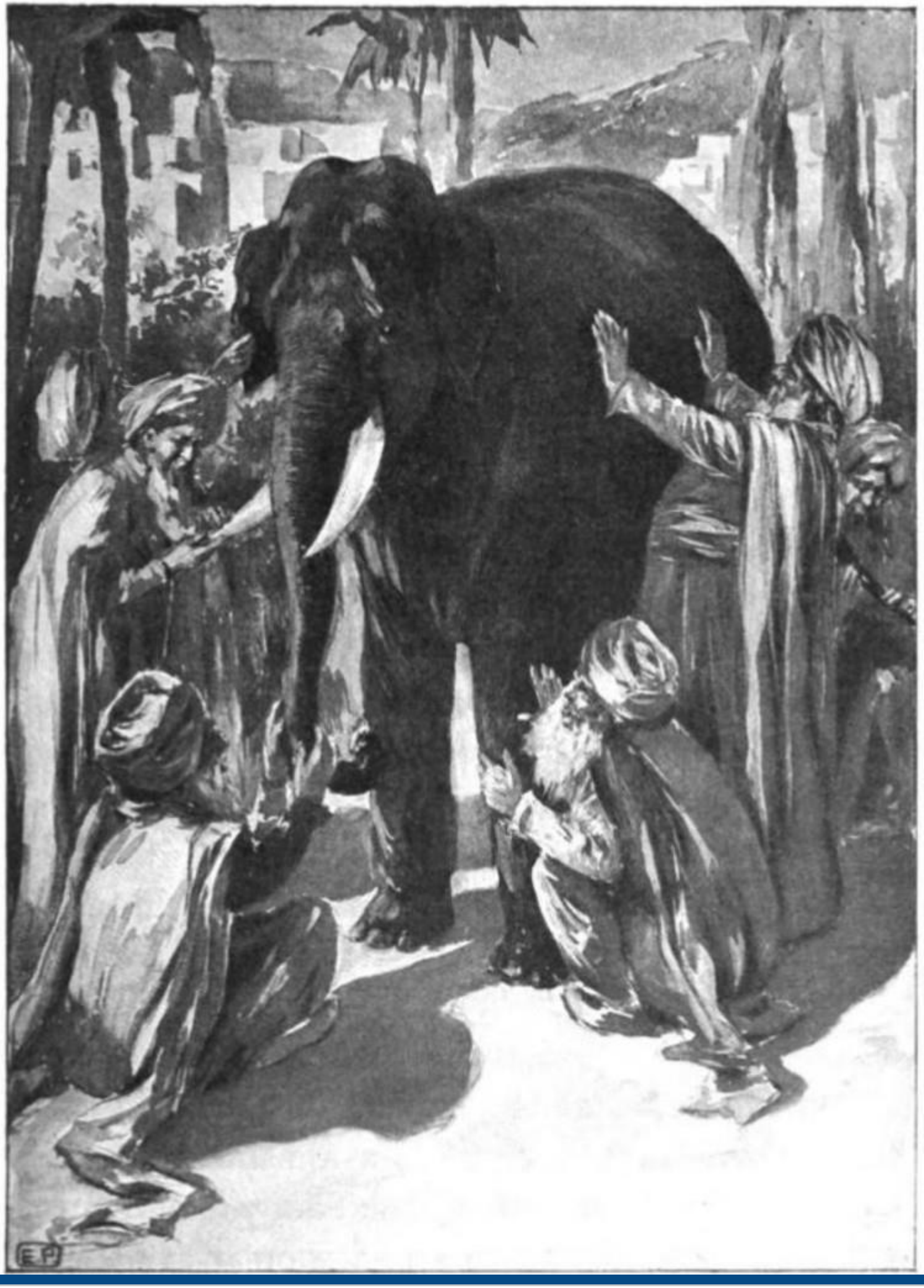
With manual vehicle

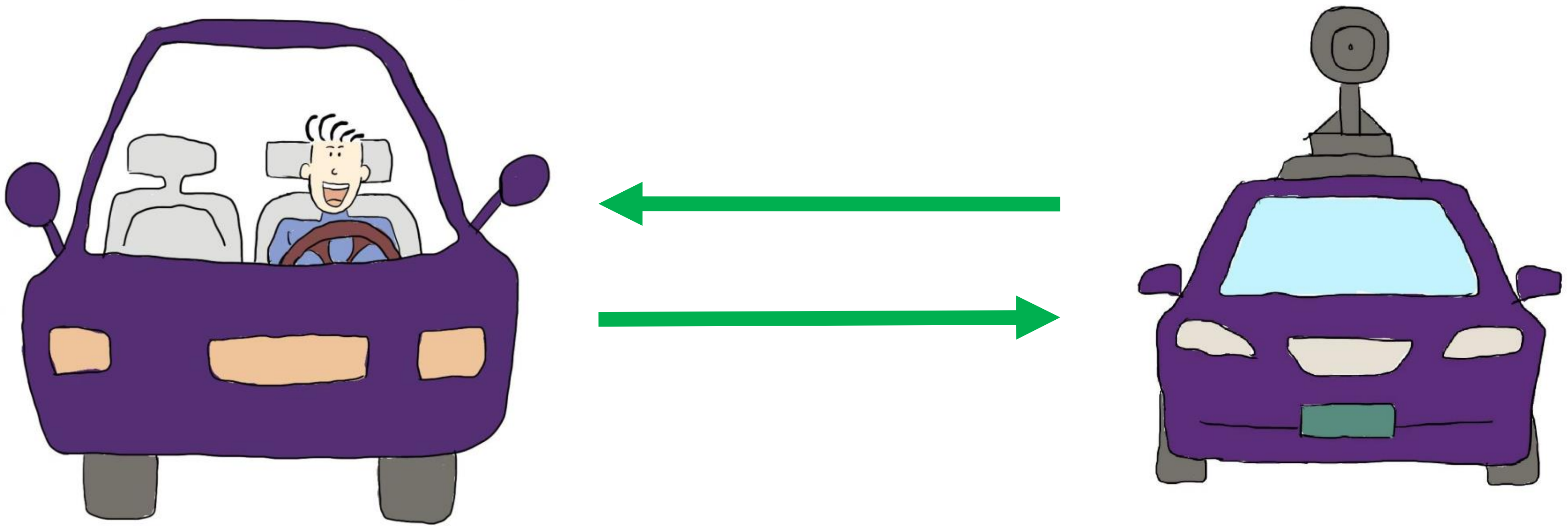


With automated vehicle

"there were no differences between reacting to AVs and MVs except for one behaviour, i.e., people are more likely to keep driving because they think they have the right of way when they were facing an AV than a MV."

Driver applied similar unwritten rules





Problematic human-AV interaction could be an emerging traffic risk.
To promote smooth interactions, AVs (developers) should know:

- how to navigate safely around human road users,
- how to interact with human road users,
- how they influence human road users,
- even, how to limit and govern human road users' risky behaviors

Thanks for listening!

Peng Liu

pengliu86@zju.edu.cn