Potentials for Enhancing People's Trust in Autonomous Vehicles

Priscilla N Y Wong
UCL Interaction Centre
University College London
ngoi.wong.13@ucl.ac.uk

Position paper submitted to CHI 2018 workshop "Interacting with Autonomous Vehicles: Learning from other Domains"

Abstract
Trust in autonomous vehicles (AVs) has been one of the most problematic issues that have been preventing them from going onto public roads. Just one accident caused by an AV can exert a significant impact on people's trust in these systems to drive safely. This paper outlines my previous research investigating people's perception of risk towards AVs and how that influences their judgements of their safety and likelihood of causing an accident. In conducting this research, I compared drivers' negotiation behaviors with AVs as opposed to other human drivers. The results of this research raised important questions about the appropriateness of our current methods for studying people's acceptance of AVs. My interests in attending this workshop are to explore the potential for overcoming people's concerns by creating better means of studying the public's perception of future events with AVs and to develop new forms of social interactions such as those occurring in the car with voice assistants and car-to-car communications.

Author Keywords
Autonomous vehicles; Driving; Trust; Acceptance

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Background

The possibility that AVs can be driven on road is highly tied to human's attitude towards AV safety. If people perceive AVs as safe, the public's higher level of trust and acceptance may speed up lawmakers' decision to legalize the operation of AVs on the road. In my previous master's dissertation [4], I conducted a questionnaire study which aimed to understand people's, in particular drivers', current state of acceptance and how their attitudes towards AVs correlated with their negotiation behaviors and their safety concerns over AVs.

Negotiation and AVs. In my previous study I first explored drivers' negotiation of limited space with other cars on road. The aim was to investigate whether or not the identity of the opposite car (i.e., a manual car or an AV) had an effect on drivers' behaviors and their willingness to negotiate spaces on road. The findings informed us that drivers' understanding of the unwritten rules (i.e., how drivers should behave towards other drivers or vehicles) did not differ, no matter if they were interacting with human drivers or with AVs.

Through observing how participants negotiate spaces with AVs, the study implicitly highlights that drivers may trust AVs and human drivers equally when making social negotiations despite the lower level of anthropomorphism in AVs. The absence of social interactions on the road (e.g., eye contact and hand gestures), which were thought to be the very essence of the 'common sense' on road [3], did not affect their behaviors, and hence acceptance of AVs. However, this might suggest the opposite: the general public actually had very little knowledge of AVs and was unable to imagine the differences between the experiences in fictitious scenarios.

Safety and AV accidents. According to the dual-processing model [1], the level of safety people perceive from AVs can be easily manipulated by the subjectivity of the information that they receive. To test this idea, I collected people's survey responses of their safety ratings on AVs after presenting them with news reports of a fictitious accident caused by an AV. Participants who read a vignette that heavily reported on factual information (e.g., how AVs have been statistically reducing accidents and fatalities), later reported AVs to be less dangerous than participants who had read a vignette that focused on subjective information (e.g., the victim's pre-accident activities). Nonetheless, participants generally felt unsafe regardless of which report they received. This suggests that with only one accident, it is enough to cause significant damage to people's trust in AVs. The greater benefit of the technology was not acknowledged and did not generate a positive influence on people's psychological perception on AVs.

The analysis of attitude scores showed that participants' pessimistic attitude towards, and high-risk concern about, technological developments were correlated with their low safety ratings. It would seem that the public is still not yet convinced by the progress that the tech industries have been making. This poses a great challenge for the future implementation of AVs.

Motivation

With increasingly more AVs on road, the likelihood of accidents involving an AV will inevitably increase, regardless of whether the cause is a human error or a

technical fault. The risk of accidents seems especially likely in these early stages of AVs being tested on the road. It is obviously important to ask the question, 'how can we minimize the number of accidents on road by using AVs?', as this is one of the fundamental goals of the development of AVs – eliminating human errors out of the driving equation. However, I believe that while current advanced technology is equipped to tackle the issue, it will not be possible to guarantee perfection or a bug-free system as AVs are also programmed by humans. Rather than perfecting the technology to settle people's uneasiness about the technology, I believe it is a two-way street – people's state of mind should be positive about AVs and people will be able to appreciate the beauty of the technology.

I believe that there is a wide range of unexplored potential that may help the public's apprehension towards AVs. My interest is to understand people's underlying psychological and social processes in relation to their trust in AVs and to develop concrete improvement strategies from the findings which may help the public overcome their concern with the safety of AVs. The challenge is to effectively identify appropriate approaches.

A major challenge for research is that the majority of the general public have not yet had expertise with AVs on the road. As a result, people's imaginations of fictitious scenarios used in evaluation studies may not be accurate. Research in this area would benefit from providing participants with stimuli, such as design fictions, to help them imagine what the future would be like with AVs. What experts are assuming to be true from their research might not turn out to be the same in real life as people are not particularly good at picturing future events about unfamiliar matters [2].

I am also interested in exploring how potential interfaces and features may be added into the system and how they may impact on people psychologically. Possible features are not limited to in-car activities (e.g., voice assistants) but also car-to-car communication and other interactions that AVs form with the outside world.

By attending this workshop, I would like to explore in what ways our understanding of novices' perception can be strengthened and how different possible features may play a part to contribute to enhancing the public's trust in AVs. I am particularly interested in exploring how underlying human attitudes correlate with social behaviors and psychological responses towards AVs, and in understanding how human's fundamental evolutionary needs can be utilized to form future improvements for AVs.

Bio

Priscilla Wong is currently an MSc student in Human-Computer Interaction at University College London. Prior to this she completed an MSc in Organization Psychology and a BSc degree in Experimental Psychology. Her previous research has investigated people's trust in AVs [4]. Her BSc dissertation research investigated how immersed people become when watching Let's Play gaming videos – the findings of this research were published at CHI Play 2017 [5].

Acknowledgement

I would like to thank my supervisor, Dr. Duncan Brumby at UCL Interaction Centre. He has provided me with immense support and guidance in shaping my research ideas.

References

1. D. Kahneman and S. Frederick, "A model of heuristic judgment," *The Cambridge handbook of thinking and reasoning*, pp. 267–293, 2005.
2. A. Salovaara, A. Oulasvirta, and G. Jacucci, "Evaluation of Prototypes and the Problem of Possible Futures," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 2064–2077.
3. C. Tennant, S. Howard, B. Franks, M. Hall, M. Bauer, and S. Stares, "The ripple effect of drivers' behaviors on the road," ThinkGoodMobility, 2015.
4. P. N. Y. Wong, J. M. Rigby, and D. P. Brumby, "Game & Watch: Are Let's Play Gaming Videos as Immersive as Playing Games?," in *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*, 2017, pp. 401–409.
5. P. N. Y. Wong, "Self-Driving Cars – Multicultural Drivers' Negotiation of Road Spaces and Safety Ratings on Accidents," *MSc Thesis, Dept. of Psychological and Behavioral Science, London School of Economics and Political Science*, 2017.