Human-machine communication in self-driving cars: verbal feedforward and feedback increase trust, but not the intention to use it. 

**Introduction**

Autonomous driving vehicles (ADVs) seem to be the primary goal of nearly all car manufacturers. More and more highly automated driving systems are implemented into newly developed cars (Dejsure & van den Brand, 2019). Technical developments in the field are fast and continuously improving. There is little doubt that self-driving cars will significantly impact our society, such as drastically decreasing greenhouse gas emissions and reducing the number of traffic-related injuries and fatalities (Benedikt & Becker, 2017; Nye 2019). However, in order for these positive aspects to be realized, it is necessary to be able to use these vehicles in public. This means public acceptance is the crucial factor for autonomous vehicles. Still, it is not precisely clear yet which factors modulate the acceptance of self-driving cars. To increase acceptance and trust, we created a virtual self-explaining car informing passengers before actions.

In this study, we investigate the current attitude of the German and Austrian population (N = 8599) in a virtual reality experiment. Participants experienced a ninety seconds virtual drive in one of three conditions: auditory feedforward and feedback, while head tracking data and a simplified gaze vector was recorded. Our main finding is that a self-explaining car does have a positive impact on trust but does not influence the intention of using such a car.

Additionally, we can show gender and age effects with female participants being generally less trusting in all conditions than male participants and a general decrease of acceptance with increasing age. Results are in line with previous human-machine research, supporting found gender differences and the fact that communications increase driving safety and negative emotional responses. These findings reveal a need for a well-balanced self-explanatory artificial intelligence to enable human-machine interactions that foster safe traffic behavior and increase trust as well as the willingness to use such technology.

**Experiment Design**

To be able to make statements about the general attitude of an entire population, a large number of subjects in a wide age spectrum are required. In the study at hand, we present a dataset of 8599 participants (3407 female) answering an adapted technology acceptance questionnaire (TAM V2). We use a setup inside the MS-Wissenschaft, a traveling exhibition ship of German’s ministry of research and education, across 28 cities in Germany and three cities in Austria over six months. The participants include 4465 male and 3386 female subjects (M = 30, SD = 3.45). However, the remaining participants identify as intersex or female with female participants being generally less trusting in all conditions than male participants. These findings were confirmed in the second setup of the experiment.

**Methods**

Fully autonomous vehicles as cognitive agents will be one of the notable use cases of artificial intelligence in the middle of modern societies. However, the potential positive effects of this technology will only apply if autonomous cars are accepted by most of the society the cars will operate in. To increase acceptance and trust, we created a virtual self-explaining car informing passengers before actions. In the study, we investigate the current attitude of the German and Austrian population (N = 8599) in a virtual reality experiment. Participants experienced a ninety seconds virtual drive in one of three conditions: auditory feedforward and feedback, while head tracking data and a simplified gaze vector was recorded. Our main finding is that a self-explaining car does have a positive impact on trust but does not influence the intention of using such a car.

**Results**

In this study, we investigate the current attitude of the German and Austrian population (N = 8599) in a virtual reality experiment. Participants experienced a ninety seconds virtual drive in one of three conditions: auditory feedforward and feedback, while head tracking data and a simplified gaze vector was recorded. Our main finding is that a self-explaining car does have a positive impact on trust but does not influence the intention of using such a car.

**Conclusion**

Although driving condition has a strong effect on the acceptance, it does not affect different aspects of acceptance the same.

- For older ages we observe an increase in overall acceptance of self-driving cars.
- As age increases we observe an initial decrease in acceptance of self-driving cars.
- Female participants scored lower on all acceptance factors compared to their male counterparts.

A self-driving car with an explainable AI, although it increases trust, it does not increase the intention of use!

**Authors**

Farbod N. Nezami1,‡, Maximilian A. Wächter1,‡, Ashima Keshava1, Hristofor Lukavon1, Marc Vidal De Palol1, Gordon Pipa1, Peter König1,‡

1 Institute of Cognitive Science, University of Osnabrück, Osnabrück, Germany
2 Department of Neurophysiology and Pathophysiology, Center of Experimental Medicine, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
‡ Shared First Authorship
* shared senior authorship
*Correspondence: Farbod N. Nezami, fnosratnezami@uos.de

**Analysis pipeline**

- Data cleanup
- OptBin Age
- MANOVA
- Linear Discriminant Analysis
- Calculating effect sizes

**Main findings**

- Strong effect of age group - for male counterparts.
- Strong effect of gender - As age increases we observe an initial decrease in acceptance of self-driving cars.
- Interaction of gender and age group - for older ages we observe an increase in overall acceptance of self-driving cars.
- Female participants scored lower on all acceptance factors compared to their male counterparts.

**Demographic factors such as age and gender has bigger impact on the acceptance of self-driving cars.**