

# **INFORMED TRUST – AN EXTERNAL USER INTERFACE FOR HIGHLY AUTOMATED VEHICLES**

**Alexander Mankowsky**  
**Stephan Mücke**  
Daimler AG  
Germany

Paper Number 19-000304

## **ABSTRACT**

Traffic research so far was focused on accidents and accident prevention. With the introduction of automated or semi-automated cars into the public realm however, the question is, how automated cars must be designed to blend in.

The presentation is based on the approach of traffic seen as a cooperative activity, where people are implicitly collaborating in the frame of given traffic rules. The Perception Action Model (PAM, Stephanie Preston 2007) in short suggests that perception and action are inseparable - people perceive actions of others and act immediately. This process is mutual and draws on empathy to predict the activities of the others. “Mind-Reading” (Eric Kandel, 2012) enables to read mood, energy, intention out of the other traffic participants.

Today it seems sufficient to perceive the way other drivers move their vehicle to trigger the Mind Reading process and enable predictive behavior. In case of automated vehicles, human perception has to be triggered properly to avoid misinterpretation or just wrong results of the empathic process.

Based on this approach the presentation introduces an experimental external user interface for highly or fully automated vehicles to address the underlying functionality of traffic. It includes the following features

- Indication of highly automated mode according to and exceeding the SAE recommendation
- Signaling of cooperative behavior
- Communication especially with vulnerable road users.

The system is demonstrated in various everyday driving situations on highways and city roads, such as merging traffic or a stop for pedestrians at a cross walk. Especially the aspects of an intuitive and intercultural understanding are discussed.

In addition, the external user interface can be used to share the vehicle’s knowledge about impending dangerous situations with its immediate surroundings – thus helping others to avoid potential accidents.

The presentation aims to demonstrate and discuss how an intuitively designed external user interface can help build informed trust in highly automated vehicles as a major factor of success and even give back to society by sharing its situation awareness.

Informed Trust is a major factor of success for Real-Word Deployment of Automated Driving Systems.

Most of the time our movement through the public realm is simply uneventful, and it is so because humans are cooperating with one another to make it so [1].

Traffic research has so far focused on accidents and accident prevention. With the introduction of automated or semi-automated cars into the public realm, however, the focus changes to understanding the magic of normal traffic and asking the following questions: how do we manage to avoid bumping into each other all the time? What will happen when automated cars are mixed into the magic? Automated cars must be designed to blend into this process; otherwise, known types of accidents will only be replaced by new types of accidents.

We have developed a conceptual framework, building on Social Sciences and Neuroscience, to understand the behavior of the involved actors in mobility. *Informed Trust* is our approach to designing self-driving cars in such a way that allows all actors to achieve agency. The overarching goal is to enable all actors in mobility to *produce a successful and harmonious mobility*, with fewer accidents than in today's traffic.

## CONCEPTUAL FRAMEWORK

### What is Mobility?

We are walking in crowded places, driving in crowded streets, and usually nothing bad happens – we manage. For the eye, it is mostly just a flow of people, cars, cycles and other means of mobility. “*The sociologist Lyn Lofland argues persuasively that the ordinary flow of movement on big - city sidewalks should be regarded as a collaborative production — a hard - won achievement in “cooperative motility” that requires the most sensitive attention to the subtle signals other pedestrians issue as to their intended course and speed.*” Lyn Lofland and others [2] are comparing our movement with a dance: we assume the movements of the others, as they do our own, and adapt continuously. As a result of the cooperative mass behavior, a rhythm emerges. Many cities are producing their own rhythm on a daily basis – you have to adapt, when switching from slow motion Charlotte, North Carolina to busy and dense NYC. For our topic, it is important to notice that walking, driving, riding a bike or a scooter is, from the standpoint of cooperation, the same. Somehow, we manage to predict the behavior of the others and adapt immediately. With self-driving cars, we introduce robots into the collective dance. Since we have not yet experienced robotic behavior in the public realm, the smooth collective dance may not happen. Practically speaking, will their behavior provoke dangerous situations? What can we do to avoid this?

### Mobility is a cooperative activity par excellence

#### The Role of Empathy in Mobility

Our approach is based on traffic seen as a cooperative activity, where people are implicitly collaborating in the frame of given traffic rules. Normally, we do not have to think while moving. We avoid accidents automatically. It is a subconscious process. One model from Neuroscience to explain our capabilities is the Perception Action Model, from Stephanie Preston and Frans de Waal [3]. In short, it suggests that perception and action are inseparable - people perceive actions of others and act immediately. This process is mutual, and draws on empathy to predict the activities of the others.

The need for empathy as a basic function of life is simple: individually, we have to find out about the intentions of other actors in our environment, and decide to run, approach, or simply freeze<sup>1</sup>. Therefore our perception scans our environment for signs of life – most importantly, for eyes watching us.

You can test it for yourself: you will be aware of the first raindrops at the window, but soon they will be not interesting anymore – they do not show signs of intention. They are not *alive*. However, if a fly approaches the same window, your perception will give you notice. Our eyes are moving more than a 100.000 times a day to achieve this, without giving us any awareness that they do so. Our perception-action mechanism runs on its own.

A hypothesis on empathy is that we incorporate the movement of other living beings. While doing so, a feeling emerges in us as we wonder how the other living being might feel and what it is up to. Neuroscientists like Erik

---

<sup>1</sup> Freezing in the face of danger implies that we are not able to drive accident free without assistance systems

Kandel [4] call this process 'Mind Reading'. It works well with living beings showing a direction when moving, like all mammals do.

Applied to traffic, it enables us to read the mood, energy, and intentions of the other traffic participants.

Today it seems sufficient to perceive the way other drivers move their vehicles to trigger the Mind Reading process and initiate predictive behavior. If empathy is impaired, like in people who suffer from autism, driving can become quite exhausting, which underscores the importance of empathy as a basic functionality for fluid cooperation in traffic.

**With self-driving cars however, empathy will fail, unless measures are taken design wise.**

### **Non-deterministic automation in Self-Driving Cars**

Self-driving cars are complicated constructs based on Machine Learning (ML) for the interpretation of the traffic scene. ML is basically a statistical process. Training data are fed into the system, until the systems can label their environment sufficiently. Since the training data and the real environment are never the same, the labeling cannot be deterministic and proven 100% right. ML based systems are therefore non-deterministic systems. Possible errors are falling into three categories: mismatch in given categories, like a mismatch between a pedestrian and a cyclist, or a total failure like labeling a golf players portrait as a golf ball, caused by golf balls visibility in too many pictures used for training the ML system. Therefore, the non-deterministic ML systems must always be framed by a classical deterministic system as a backstop. In self-driving cars, for example, a simple radar system, which overrides the ML system and stops the car, when something is sensed on the road that contradicts the traffic scene created by the ML program.

The resulting behavior of a self-driving vehicle can be weird [5] for us, because our prediction skills fail against non-deterministic automation.

**To enable our predictive capabilities, the car needs to provide hints for our perception, so that we can 'feel' it the way we can feel other participants in traffic.**

### **Design Approach - Informed Trust**

Informed Trust is opposed to blind trust: the goal of *Informed Trust* is to enable all the participants in a traffic scene to act: in short, to foster agency. Without agency, some actors in traffic will end up as victims by design, others as villains, also by design<sup>2</sup> [6].

To enable agency when dealing with self-driving cars, we have to prepare design elements working as handles for empathy, design elements functioning as enablers for learning, and finally, define procedures for when the systems degrade.

#### **Informed Trust: Empathy and insight**

Empathy: Feel the car

- Has it seen me? Is it friendly? Is it coming?
- Which direction is it heading? Is it relaxed or strained?
- Shall I help? Is it helping? I miss it already!
- Is it aware of me?

Insight: Learning process

- Display the mode of the car: is it in self-driving mode or not

---

<sup>2</sup> Compare 'Moral Machines' from MIT, Iyad Rahwan, Jean-Francois Bonnefon, Azim Shariff. Once you enter the game-like traffic situation, you end up as a executioner by design. <http://moralmachine.mit.edu/>

- Avoid misinterpretation of the role of a person sitting in the driver's seat
- Predictability of the car in different situations
- Offer simple conceptual models of the technical functions

Lessons: Graceful degradation

- Clear functional steps in case of break down
- Universally understandable warning signs to driver and traffic
- Standard procedure in case of accident detection

**Empathy** is based on functions of our perception. Our perception is scanning our environment continuously for body-movement and eyes. Once we encounter life, the prediction process starts. For a self-driving car, the communication of intention and state through movement can be implemented in several ways. The body of the car itself can be set in motion, parts of the cars body can be used, the way it is moving ahead, creeping, accelerating, decelerating and so on. The body of the car can be equipped to simulate muscle movement or wrinkles.

We experimented successfully with a wake-up motion of the car similar to an animal<sup>3</sup> [7]. The result was awareness on the side of pedestrians in front of the car when it 'awakes' in the parking lot, without any alarmism. The effect triggered a reaction which was felt by test participants as quite natural.

The importance of eye contact is paramount for informed trust. It can be implemented through various means, such as an LED pointing at people within a critical distance. These LEDs can be programmed to follow somebody, communicating to them a 'Hey it's you'. A short blink is similar to what we are doing regularly when walking the street. We also used a virtual shadow on the surface of the concept-car Vision URBANETIC, which is more complicated to implement, but very natural. Once we get used to such a shadow, we get alarmed when it is missing, and are so able to avoid a possible accident. Moreover, directed sounds are helpful to underscore urgency to make contact, especially for the visually impaired.

To ensure the empathic effect, the tools should work together organically. Our perception is watching for signs of life, so we should strive for a 'half-life' approach, as we know it from animated objects.

**Insight** needs simple and dependable signaling on the side of the self-driving car. In traffic school for children, it should be as easy as possible to teach them how to interact with such a car.

Teaching example for children:

- 'Is the car in self-driving mode?' - Watch for the signaling lights
- 'Has it seen you?' - Check signs of eye contact
- 'No signs of eye contact?' - Go away

To achieve this scenario, the car must show its state as self-driving in 360 degrees around the car. The signaling should be arranged in such a way that the signal is applied in direction of the driver's seat, as it is the ritualized direction of the short gaze. The mode signaling has to be standardized. Standardization here is seen as the single most important success factor for the application of the technology.

**Lessons** have to be prepared for the case of a system failure. When the system is failing, it has to use a standard procedure, also known as Graceful Degradation in automation. To know this procedure will be most important for

---

<sup>3</sup> Clifford Nass, Psychology of Automated Vehicles, Stanford, 2013.

Nass suggests to compare self-driving cars with domesticated animals and design the interaction-narrative accordingly. <https://www.youtube.com/watch?v=hrxf7lG-j9c>

people inside and outside of the car. Graceful Degradation is fundamental to the debate on ethics and automation. From the perspective of the program running a self-driving car, an accident is always a failure out of missing or incorrect data, with failed prediction as the outcome. The concept of Graceful Degradation can be helpful here, defining the standard procedure in such a case. Straight emergency braking will give the other actors in the accident scene the opportunity to act meaningfully, once they know for sure that all self-driving cars will behave in exactly this way.

## INTRODUCING AN EXTERNAL USER INTERFACE FOR AUTOMATED VEHICLES

### Sensor setup, high-definition maps and digital integration

The environmental perception system of automated vehicles consists of sensors and perception algorithms. Their customized sensor setup uses various sensor types like radars, lidars, cameras and ultrasonic sensors as well as microphones. It provides a robust 360-degree field-of-view around the vehicle to enable the vehicle to handle all relevant use cases and maneuvers [1].

Deep Learning algorithms allow an object classification. The example in figure 1 illustrates different object classes by color-coding: vehicles are shown in blue, pedestrians in red, the road in purple, and traffic signs in yellow. It even includes objects partially blocked from the automated vehicle's perception system, such as pedestrians partially hidden by other vehicles.

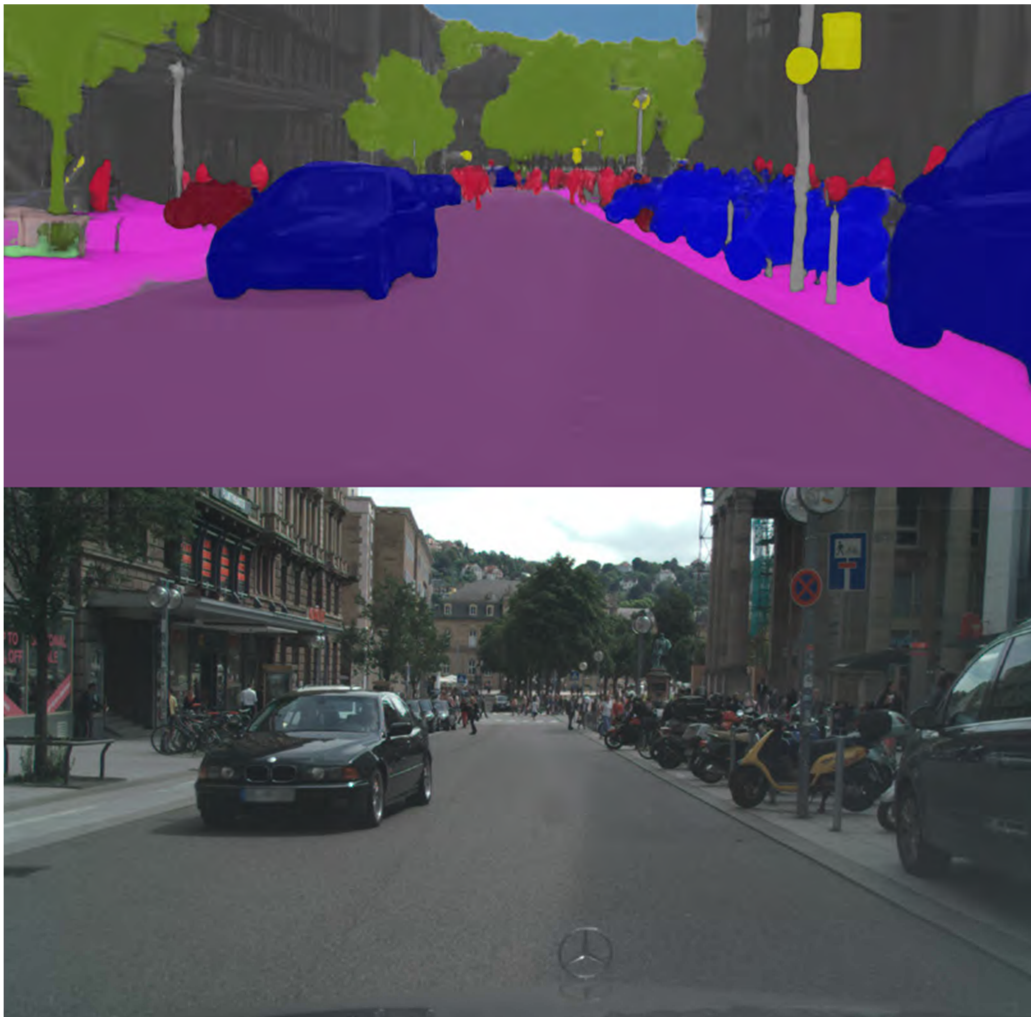
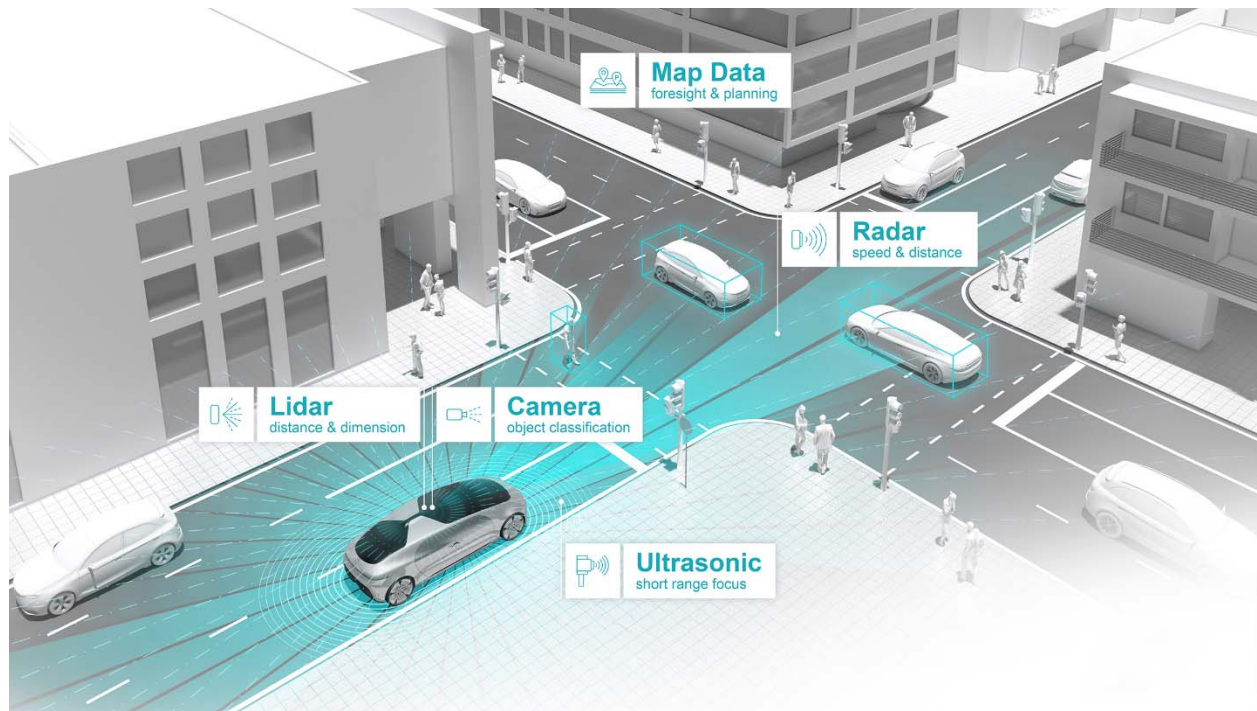


Figure 1. Object classification based on camera images [8]

A high-definition digital map defines the operational design domain and allows locating the automated vehicle and surrounding vehicles, pedestrians and objects. Environment modeling anticipates and predicts what other objects on the roadway might do.

Automated vehicles will be integrated within a digital infrastructure, which is capable of connecting to a Fleet Operations Center, data sources from public agencies (e.g. traffic signal data from local road authorities) and data from other vehicles. It also has an interface to the cloud where information such as status reports, relevant traffic, weather, incident and construction zones can be accessed.



**Figure 2. Sensor setup of an automated vehicle**

### **External User Interface**

#### **Indication of automated driving mode**

As mentioned in the context of interaction with pedestrians, the indication of automated driving mode is essential for the acceptance of automated vehicles. It allows pedestrians and other road users to identify automated vehicles, learn to read their intentions and thus build the aforementioned “informed” trust.

In 2018, Mercedes-Benz and Bosch started an Automated Valet Parking Service as a pilot project in Stuttgart [9]. Figure 4 shows a first realization to indicate the automated driving mode using turquoise illumination of already existing lights in the exterior mirrors and the third brake light.



**Figure 3. Automated Valet Parking: Indication of automated driving mode [9]**

Turquoise was chosen as indication of automated driving mode because it is unique, even peripherally visible and not used for any other exterior vehicle lighting. An internal Mercedes-Benz study with pedestrians revealed that a majority of the participants preferred turquoise and all of them voted for the 360° indication. Mercedes-Benz is supporting activities of the SAE, to develop an exterior lighting concept for automated vehicles.

#### **Intuitive communication with other road users**

In many situations road users, especially pedestrians should know instantaneously and reliably what an automated vehicle will do next.



**Figure 4. Communication with a pedestrian wanting to cross the road [10]**

### **Signaling cooperative behavior**

An important goal for automated vehicles is to reduce the likelihood of accidents and to mitigate risks for passengers and other road users by operating like a defensive and attentive driver who consistently monitors the driving environment and responds appropriately and safely to changing conditions. However, trust in automated vehicles doesn't only require their cooperative behavior. They also must inform their immediate surroundings about their intentions in a way that can be understood intuitively.

Again, the combination of an indication of autonomous driving mode and an intuitively understandable message help other drivers to learn the behavior of automated vehicles and to build "informed" trust.

### **Warnings**

In addition, the external user interface is used to share the automated vehicle's knowledge about impending dangerous situations with its immediate surroundings – thus helping others to avoid potential accidents.

Especially vulnerable road users like pedestrians and bicyclist who are easily missed or obscured and do not have their own assistance systems may profit from this feature.

**Local Hazard warnings:** Driving safely and smoothly in complex environments requires detailed and ongoing awareness of the real-time traffic situation, coupled with the ability to forecast future traffic developments.

## **CONCLUSIONS**

The presentation aims to demonstrate and discuss how an intuitively designed external user interface can help to build informed trust in automated vehicles and to create an additional safety benefit for others by sharing its situation awareness. The Mercedes-Benz Experimental Safety Vehicle 2019 (Experimentales Sicherheitsfahrzeug ESF 2019) will show how these considerations can be implemented in future vehicles.



## REFERENCES

- [1] Lofland, Lyn H. The Public Realm, p. 29. Taylor and Francis, 1998. Kindle Edition.
- [2] Greenfield, Adam. Radical Technologies: The Design of Everyday Life. Verso, 2017. Kindle Edition, Location 1296
- [3] Preston, Stephanie D. and de Waal, Frans B.M. (2001) Empathy: Its ultimate and proximate bases. [https://greatergood.berkeley.edu/images/uploads/Preston\\_dewaal2002.pdf](https://greatergood.berkeley.edu/images/uploads/Preston_dewaal2002.pdf)
- [4] Kandel, Eric: Age of Insight, 2012. The hypothesis may explain why some people are afraid when encountering spiders or snakes. We cannot replay their behavior, and therefore not predict what they are about to do. Could be similar when encountering robots and self-driving cars
- [5] The State of Self-Driving Cars in 2018, <https://www.youtube.com/watch?v=1Jf1ZM-ho4o>
- [6] Rahwan, Iyad; Bonnefon, Jean-Francois and Shariff, Azim: Moral Machines. MIT. <http://moralmachine.mit.edu/>
- [7] Nass, Clifford: Psychology of Automated Vehicles, Stanford, 2013.
- [8] “Reinventing Safety: A Joint Approach to Automated Driving Systems”, Mercedes Benz in cooperation with BOSCH, VSSA, 2018 Mercedes-Benz Research & Development North America, Inc. and Robert Bosch LLC. <https://www.daimler.com/dokumente/innovation/sonstiges/vssa-mercedes-benz-and-bosch.pdf>
- [9] Future insight: <https://media.daimler.com/marsMediaSite/de/instance/ko/Mercedes-Benz-FutureInsight-Human-first-Empathie-als-Anker-in-der-digitalen-Transformation.xhtml?oid=41918249>
- [10] Automated Valet Parking: [https://media.daimler.com/marsMediaSite/de/instance/ko.xhtml?oid=23076118&ls=L3NIYXJjaHJlc3VsdC9zZWFiY2hyZXN1bHQeGh0bWw\\_c2VhcmNoU3RyaW5nPXZhbGV0K3Bhcmtpbmcmc2VhcmNoSWQ9MCZzZWFiY2hUeXBIPWRldGFpbGVkJmJvcmlcnM9dHJ1ZSZyZXN1bHRJbWZvVHlwZUIkPTQwNjI2JnZpZXdUeXBIPWxpc3Qmc29ydERlZmluaXRpb249UFVCTEITSEVEX0FULTImdGh1bWJTY2FsZUluZGV4PTAmcm93Q291bnRzSW5kZXg9NQ!!&rs=1](https://media.daimler.com/marsMediaSite/de/instance/ko.xhtml?oid=23076118&ls=L3NIYXJjaHJlc3VsdC9zZWFiY2hyZXN1bHQeGh0bWw_c2VhcmNoU3RyaW5nPXZhbGV0K3Bhcmtpbmcmc2VhcmNoSWQ9MCZzZWFiY2hUeXBIPWRldGFpbGVkJmJvcmlcnM9dHJ1ZSZyZXN1bHRJbWZvVHlwZUIkPTQwNjI2JnZpZXdUeXBIPWxpc3Qmc29ydERlZmluaXRpb249UFVCTEITSEVEX0FULTImdGh1bWJTY2FsZUluZGV4PTAmcm93Q291bnRzSW5kZXg9NQ!!&rs=1)