Elaborating Feedback Strategies for Maintaining Automation in Highly Automated Driving

Philipp Hock[⋆‡], **Johannes Kraus**[‡], **Marcel Walch**[⋆], **Nina Lang**[‡], **Martin Baumann**[‡]
[⋆]Institute of Media Informatics
[‡]Institute of Psychology and Education, Dept. Human Factors
Ulm University, Germany
[philipp.hock, johannes.kraus, marcel.walch, nina.lang, martin.baumann]@uni-ulm.de

ABSTRACT

Human errors are a major reason for traffic accidents. One of the aims of the introduction of automated driving functions in vehicles is to prevent such accidents as such systems are supposed to be more reliable, react faster with higher precision. Therefore, we assume that an increase of automation features will also increase safety. However, when drivers are not willing to relinquish control to the vehicle, safety benefits of automated vehicles do not take effect. Therefore, convincing drivers to actively make use of the automation when appropriate can increase traffic safety. In this paper we investigate the influence of system feedback in proactive, safety critical takeover situations in automated driving. In contrast to handover, which is initiated by the system, proactive takeover is initiated by the driver, who's intention for steering the car is the reason for driving manually. We compare auditory feedback with audio-visual feedback realized as a virtual co-driver in a user study. We conducted a virtual reality simulator study (n=38) to investigate how system feedback influences the willingness of drivers to relinquish control to the vehicle. There were three conditions of system feedback: in condition *none* no feedback was given, in condition *audio* spoken feedback was given, and in condition *co-driver* additionally to audio feedback, a virtual co-driver on the front passenger seat was displayed. Our research provides evidence that system feedback can lead to an increase of willingness to maintain automation and to follow its safety related advices.

ACM Classification Keywords

H.1.2 User/Machine Systems:: Human factors; H.5.2 User Interfaces:: Natural language, Prototyping, User-centered design

Author Keywords

Automated Driving; Persuasion; Avatar; Affective HCI, Human-robot interaction; Anthropomorphism

DOI:http://dx.doi.org/10.1145/3003715.3005414

INTRODUCTION

Research on automated driving vehicles has two approaches: the first one aims at introducing a fully autonomous car that does not result from a stepwise approach of reducing human operator control, like the *Google Self-Driving Car Project* [6], the other approach is an iterative improvement of Advanced Driver Assistance Systems (ADAS) up to fully autonomous cars [22]. Both approaches coexist at the moment. The latter includes a transition period of semi-automated driving, where the driver can take control back from the automation and decides when and under which circumstances he wants to drive manually and when to drive automatically.

In both approaches the vision of increased traffic safety achieved with (semi-)automated vehicles is one of the major arguments for these developments. The basis for these arguments are results of investigations that indicate that human errors account for over 90% of all traffic accidents [21]. And these errors could be avoided with (semi-)automated vehicles. Current automated driving systems are not completely reliable but we assume that with an increasing maturity of this technology, the use of automated driving systems will lead to a decrease of severe accidents. Besides the technological barriers that have to be solved before such vehicles can be introduced the drivers' willingness and acceptance of automated vehicles are as important. The drivers' joy of driving and the need for being in control of the vehicle are factors that could restrain the use of automated driving functionalities. In a survey with 5000 participants [15], it was shown that manual driving was the most enjoyable mode of driving.

Additionally, a lack of trust and acceptance can cause denial of automated driving. According to a study from 2014, only 44% of U.S. car drivers can imagine to buy a fully autonomous car and only 22% of U.S. car drivers can imagine to buy a partially automated driving car [29, 30]. The lack of acceptance in automated vehicles and ADAS is also shown in other literature [11, 26].

This means that the possible safety benefit of automated driving vehicles can only be fully exploited if drivers are willing to use the automated driving functions of their vehicles and consequently to give or leave the control of the driving task to their vehicle.

We conducted a study, in which we investigated an approach to convince drivers to maintain automation in safety critical situations. We investigated 3 different automation feedback conditions: (1) without feedback, (2) by communicating the system state by a computer generated voice, and (3) by communicating the system state by a projection of a humanoid avatar as co-driver. Our main contribution is to show that appropriate system feedback can increase the willingness to leave control over the driving task with the automation in safety critical situations.

In the following, we first define proactive takeover situations, then the role of trust in automation and how system transparency can increase trust is pointed out. Subsequently, an overview of persuasion, especially the power of credibility and social actors regarding persuasion is being discussed. Next, related work regarding avatars, passengers, emotions and affection is presented. In a user study, we show that system feedback has a significant influence on persuading the behavior of maintaining automated driving functionalities to increase safety on the road.

PROACTIVE TAKEOVER SITUATIONS

In contrast to handover, which is initiated by the system in most cases, takeover is initiated by the driver. It can be assumed that a driver who is willing to takeover control of the vehicle, is aware of the situation and not affected by the *out of the loop performance problem* [3]. It states when a systems is used over a longer period in an automated mode, users will encounter problems when taking back control from the system. We define proactive takeover situations as action intended by the driver to takeover control of the vehicle when driving autonomously or semi-automated. In our study, this takeover intention does not result from approaching a system boundary. This means the system is able to handle the current situation perfectly but the driver wants to takeover control due to some other reasons. Reasons for such a takeover might be boredom, impatience, dislike of the automation's driving style or distrust in the automation [17]. In this study, we investigated situations, in which the automation's driving style could be perceived as unreasonably cautious. As an example we implemented a situation where the automation decides to follow a slow lead vehicle due to safety reasons. Such a situation could result in the driver getting frustrated, which in turn could increase the driver's willingness to takeover control and switch to manual driving followed by an overtaking attempt thereby risking a severe accident. The question is how such behavior could be reduced or prevented.

RELATED WORK

This section discusses the connection between trust and feedback. Subsequently, the relation between persuasion and feedback is examined. Finally, prevailing research regarding persuasive technologies and affection, avatars and the influence of passengers is presented.

Trust

For the domain of automated driving systems, trust in automation plays an important role. Lee and See state that the lack of trust (distrust) can lead to a denial of usage of a system. When trust exceeds system capabilities (overtrust) and system boundaries are reached, the driver may not be capable of performing an appropriate action [17]. In our scenario, system boundaries are never reached, therefore only distrust is considered as an important aspect in the study.

Hoff and Bashir [8] have outlined the relationship between trust in computer systems and interpersonal trust as follows: trust in computer systems depends on performance, process or purpose of the automation whereas trust in humans rely on social factors, like ability and integrity [16]. Madhavan and Wiegmann [18] state that the progression of interpersonal trust differs from human-automation trust. Holmes, and Zanna [25] state that interpersonal trust – which is the trust between people – initially depends on the predictability of a person's action. This changes in dependability and integrity over time and in the final stage, trust is based on benevolence or faith. Human-automation trust progresses the reverse order where people assume that machines behave perfectly, causing trust based on faith. On behavior that seems erroneous, trust dissolves. Faith is replaced by dependability and predictability as primary basis of trust [17]. This means that clarifying potential errors in automation by providing system transparency can increase trust, which in turn increases usage of the automation.

Persuasion

Persuasion is defined as the act of causing people to do or believe something [1]. Fogg [5] defines persuasive technology as technology that is designed to influence people in their actions or believes. He states that an important attribute of persuasive technologies is credibility. A system that is credible, can also be persuasive whereas a system with low credibility has low persuasive power. Lee and See [17] state that computer systems have high initial credibility, according to Fogg, credibility can decline when system behavior is not comprehensible [5]. Therefore, similar to trust, system transparency and system feedback can lead to more credibility, which in turn leads to persuasion.

A lot of research has been conducted in persuasive technology in the automotive context. However, a lot of research focuses on eco-driving feedback technologies, such as [19, 14]. Examples of persuasive technologies to increase safety are numerous. For example warning sounds on seat belts, speed trailers, info panels with statistics of accidents. Research on different persuasive technologies that increases safety, however seems rare.

Another persuasive element in technology is the role of social actors [5]. Fogg proposes five types of primary cues that cause people to see technology as a social actor: physical (appearance), psychological (e.g. personality), language, social dynamics (e.g. answering to questions), social roles (e.g. a teammate or a doctor). Computer systems that are intended to take such role, should therefore entail at leas on of these features. According to the *Media Equation Theory* [24], people tend to treat computer systems as if they were real persons. Therefore, the influence of virtual avatars could lead to a similar influence as real people. This leads to the assumption that social actors providing system feedback can influence people's behavior.

Influence of Passengers and Avatars

The influences of passengers on driving behavior has been investigated in several studies so far. Fleiter et al. show that people tend to drive riskier when driving alone [4]. Additionally, in a study with teenage drivers, Simons-Morton et al. [28] found a moderation of the effect of passenger presence by type of passenger. They found that risky driving was decreased when driving with adult drivers compared to driving alone, but they found an opposite effect when the passengers were friends of the same age. Other research also shows that the influence of passengers regarding driving performance is high [2]. Above this, it is shown that collaboration between driver and co-driver leads to an increase of safety [7].

Also in regard to robotic avatars, there has been some research recently. Williams and Breazeal presented a system called AIDA (Affective Intelligent Driving Agent) [33, 35, 34], a social robot as interaction device on vehicle dashboards. In NAMIDA [10], a similar social robot setup was presented. In contrast to AIDA, NAMIDA has three grouped avatars that are able to interact with each other. These works focus on reducing cognitive load when interacting with the vehicle's infotainment system and hereby build on social properties to influence the driver and to increase trust and credibility in the automation. In [36], a companion robot as car interface is portrayed and implemented as a tablet application. In their study, affective interaction showed to be more effective in regard to informative interaction. Affective interaction means that communication is achieved through an avatar based interface whereas informative interaction means communicating facts only.

Emotions and Affection

The role of human perception and expression of affective states and emotions is crucial for social interaction. The same psychological mechanisms play a role in human-computer-interaction [23] and thus play a major rule in the endeavor of influencing drivers' behaviors. Nass et al. [20] show that social cues like emotions influence driving performance. In their study, emotional matchings between the car's audio feedback and the driver's affective state increased safety. Furthermore, it was shown that anthropomorphism can enhance trust in an automated vehicle [32]. Koo et al. show that in semi-autonomous driving, the kind of provided system feedback influences driving performance. Informing drivers about *what* and *why* something is going on resulted in the best driving performance [13]. This insight was respected when designing system feedback in our study.

EXPERIMENT DESIGN

Our study is based on the hypothesis that system feedback provides transparency and increases trust and acceptance in the automation, leading to our first research question: does system feedback has an influence on trust and acceptance of the automation. The need to understand why an automated car performs specific actions, leads to our second research question: does system feedback impact the user's willingness to maintain or abort the automation. System feedback is given in form of auditory feedback in the *audio* condition. In the *co-driver* condition, additionally to the auditory feedback, a virtual co-driver is rendered into the scene. Our second assumption is that a virtual co-driver increases the level of trust and therefore leads to higher compliance and persuasion. Accordingly, our third research question: does a virtual co-driver have a greater effect on maintaining the automation than auditory feedback.

The following hypothesis were derived from the aforementioned current state of the research literature: **H1a:** Trust and acceptance is highest in the *co-driver* condition. **H1b:** Trust and acceptance is higher in the *audio* condition than in the *none* condition. **H2a:** In the *audio* condition the automation is maintained longer than in the *none* condition. **H2b:** In the *co-driver* condition the automation is maintained longer than in the *none* condition. **H2c:** In the *co-driver* condition, the automation is maintained longer than in the *audio* condition.

We assume that takeover control of the vehicle would most likely occur when the participant approaches a slowly driving car ahead and the automation does not overtake the car by itself. The system's reason for that was a limited visual range which made overtaking the car a risky behavior. In the study, participants had to drive in an automated car on a rural road. After a short period, a slowly driving car ahead approached, where participants could cancel the automation to overtake the car. We further assumed that the type of feedback had an influence on the probability of canceling the automation.

Method

We recruited 38 participants (25 female) through flyers and mailing lists. Participants were 24 years old on average ($SD = 3.54$) and in possession of a valid driver's license since 6 years on average ($SD = 3.00$). Depending on their performance, 6-9 euros were paid as reward. The study lasted about one hour. A between subject design with randomly assigned participants was used in the study.

At the beginning, a demographic questionnaire was filled out and an introduction to the study procedure was provided in written text. The study setting was conducted in a virtual reality driving simulator by using an Oculus Rift DK2 and a Logitech G27 steering wheel. The participants additionally wore headphones to isolate them completely from reality and to increase the immersive effect of the VR setting.

To familiarize participants with the VR device and the study environment, three test scenarios were executed before the actual experiment, to ensure that test subjects were familiar with overtaking cars in the setup. This included steering the vehicle, switching between manual operation and automated driving of the car and overtaking a car ahead.

Figure 1 shows the view from the interior of the driving car in the study (left). Fog limits the visual range up to 200m and a car in front forces the automation to drive at 60km/h. On the right, the person wearing an Oculus Rift DK2 driving in the study setup.

In the scenario participants were seated in an automated car on a two lane rural road. The track was a circuit with 3.400m in diameter. Participants drove counterclockwise constantly driving a slight left turn, which was chosen because participants

time passed (s)	visual range (m)
0s	200m
120s	500m
160s	800m
200s	1100m
240s	*End*

Table 1. The Table shows the connection between time passed since the beginning of the drive in seconds and visual range at this moment in meter.

Figure 1. Left: a car approaching out of the fog (visual range: 200m) ahead of the driving car. Right: virtual reality study setup using Oculus Rift DK2.

should easily spot contraflow without deactivating the automation to swerve. The car drove autonomously with 100 km/h. During the experiment, fog was used to limit the participant's visual range to an exact amount (200m in the beginning). The fog changed after predefined periods during the study. We used fog as control variable for visual range. Participants were told that the human visual range matches the vehicle's sensor range. After 15 seconds, a car on the opposite lane was approaching to show the participants that oncoming traffic exists in the scenario. After 25 seconds, a car driving with 60km/h appeared out of the fog driving ahead. After 37 seconds, the car ahead was reached and the automation slowed down, following the car ahead.

After 45 seconds, the system gave feedback why it is not overtaking the car ahead. 40 seconds later, the system gave feedback again, the fog did not change at this point. After that, the fog changed every 40 seconds to the following visual ranges: 500m, 800m, 1100m. The relationship between visual range (vr) and time passed (tp) in the study is shown in Table 1. A constantly extending visual range instead of a random order was used, to gradually reduce the risk potential. In combination with possible growing impatience, we expected a constantly increasing demand of stopping the automation and overtaking the car ahead. Our intention was to find the point in time where participants aborted the automation. We further assumed that feedback leads to a higher degree of persuasion, which in turn leads to a longer maintenance of the automation, resulting in a later point of abortion. We decided to use a gradually changing visual range to have clearly separated situations where system feedback could be applied reasonably. This way a relation between time until abortion and persuasive effect was given.

In each change, the system gave feedback why it is not going to overtake the car ahead, e.g. "the fog is still too dense therefore I cannot overtake the car ahead." Reasons were formulated vaguely, no calculation or exact visual ranges were communicated.

Three conditions were used for the evaluation. In the *none* group no feedback was given. In the *audio* condition, the system gave auditory feedback regarding the current condition. The feedback contained information *that* and *why* the automation does not overtake the slow car ahead. In the *co-driver* condition, the same audio feedback as in the *audio* condition was presented and in addition, a virtual co-driver was sitting on the front passenger seat. The co-driver is described in the implementation section.

At the beginning of the experimental trial, in the *audio* and *co-driver* condition, the feedback system introduced itself, telling the participants that the system represents the functionalities of the vehicle and that it will deliver relevant information to the driver. In the *co-driver* condition, participants were notified about the presence of the virtual co-driver because the limited field of view of the VR device prevented spotting the co-driver when looking straight ahead.

The drive took 4 minutes in total, no matter how fast participants were driving or if overtaking the car ahead occurred. To give them an incentive to overtake the car ahead, we gave participants 1 euros for each kilometer driven. When the car drove the entire track autonomously, an amount of 6 euros were paid out. If the participants, overtook the car at the earliest point in time possible, 9 euros were paid out. Participants were told that a manual takeover that results in an accident or speeding, only the minimal amount of reward would be paid out. By this an incentive to drive as many kilometers as possible and comfortable without breaking the law or causing an accident in the given amount of time was created. All participants were informed about this possible bonus reward before the beginning of the drive.

After the trial, participants answered a questionnaire containing scales for simulator sickness [12], trust [9], acceptance [31] and immersion [27]. In the *co-driver* condition, appearance of the avatar regarding uncanniness was also elicited. We also asked participants how useful the given feedback was. Simulator sickness was measured before and after the simulation.

To summarize, we conducted a between-subjects study with 38 participants in a VR driving simulator. The between-subjects factor was the kind of feedback given (*none*, *audio*, and *co-driver*). As dependent variable, we measured the time until participants aborted the automation. The visual range was realized with different levels of fog. We expected that most participants would overtake the car at some point during the study because the visual range increased due to a gradually decreasing fog level as the study went on.

Implementation

We implemented the study setup with Unity 3D. The scene was presented in virtual reality because being inside a vehicle while sitting beside a virtual co-driver was essential in the study. Due to the VR setting, only a driving wheel with pedals was needed to create an appropriate amount of immersion. A Logitech G27 Racing wheel with corresponding pedals were used in the study. 3D sound was designed to trigger the perception that the origin of the voice was in the seat so that the co-driver could be identified as the source of sound. We designed the co-driver to appear as a projection (Figure 2) in order to clarify that the co-driver should not be recognized as a human being but instead as an artificial humanoid that represents the computer system and the automation. The setup with the virtual co-driver first might seem unrealistic because projecting a person onto the passenger currently cannot be realized. However, the intention of this study was to evaluate the concept of a virtual character being present in the car. A technical implementation in reality could be achieved by displaying the co-driver on the window, into a monitor on the dashboard or as dedicated hardware element somewhere in the car. We decided to use an avatar on the front passenger seat because we assumed that a high similarity with an actual human co-driver would have the biggest impact regarding persuasion.

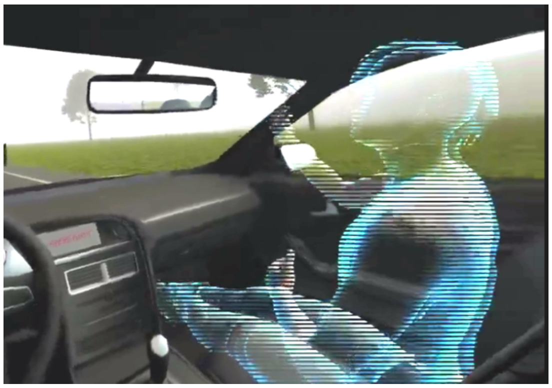

Figure 2. The avatar in the study is designed as a projection to avoid confusion whether the avatar should represent a human being or a computer system. The avatar gesticulates to increase humanoid appearance.

The voice in both conditions (*audio*, *co-driver*) was the same and was generated by a text-to-speech engine. An Oculus Rift DK2 was used as VR device. To make the character appear more realistic, we captured the avatar's motions with an Optitrack motion capturing system. In the simulation, the car had a linear acceleration of $1.6\frac{m}{s^2}$, accelerating from 60 km/h to 100km/h in about seven seconds.

System feedback

Auditory system feedback regarding the visual range occurred 5 times during the study. In the first situation, the visual range amounted 200m and the following text was spoken after 45 seconds: "The fog is too dense, therefore the automation will not overtake the car." After about 90 seconds, visual range was still 200m and the following feedback was spoken: "The fog is still too dense, therefore no overtaking maneuver will be initiated." After the fog extended visual range up to 500m (2.5 minutes), the following message was provided by the system voice: "The fog has in fact declined, overtaking however is still not safe." After 800m visual range (160s): "Sight has clearly improved, however, without risk overtaking is not possible." After 1100m (3.3 minutes): "Sight is barely restricted, but there is a certain risk potential, therefore I will not overtake the car ahead." On the last stage, where visual range amounted to 1100m, the automation still did not overtake the car while giving the feedback that the situation bears a certain risk. The feedback was chosen to be of informative and non-imperative nature, no instruction for initiating any kind of action was provided by the system voice.

Overtaking

The visual range to overtake a car ahead is calculated by the sum of the distance driven to overtake the car plus the distance driven by a potential approaching contraflow. The distance to overtake a car ahead is calculated by the sum of the following distances: the safety distance ahead, the length of the car ahead, the distance driven by the car ahead during the overtaking maneuver and the safety distance between the overtaking car and the car that has been overtaken. Pretests in the simulator show that the process of overtaking the car ahead takes about 12 seconds from the decision to take over until the overtaking maneuver is finished and the car is stable on the correct lane. If the whole overtaking maneuver takes about 12 seconds, the car travels about 300m during the overtaking maneuver. A contraflow car approaching with $100km/h$ travels about $335m$ in 12 seconds. This makes a distance of $300m + 335m = 635m$ and therefore the minimum visual range to overtake a car ahead.

Results

The abortion time of the automation was analyzed in the three conditions *none* where no feedback was given, *audio* where auditory feedback was given and *co-driver* where additionally to the audio feedback, a virtual co-driver was displayed on the front passenger seat, acting as proxy for the automation. Participants in the *none* condition aborted the automation on average after 68.8 seconds ($SD = 30.1s$), 154.8 seconds in the *audio* condition ($SD = 65.0s$) and after 136.4 seconds ($SD = 67.0s$) in the *co-driver* condition (Figure 3). Table 2 shows the percentage of participants who aborted the automation according to the condition and the visual range.

	visual range				
Condition	200m	500m	800m	1100m	no abort
none	92%	8%	0%	0%	0%
audio	29%	36%	7%	0%	28%
co-driver	50%	8%	17%	17%	8%

Table 2. The Table shows at which visual range participants aborted the automation in relation to the conditions with no feedback (none), auditory feedback (audio) and feedback with co-driver (co-driver).

Between the *audio* condition and the *none* condition lies on average 86 seconds. As for the abortion times the precondition of normal distribution was violated (Anderson-Darling Test; $p < 0.05$), a non-parametric alternative to ANOVA had to be applied. A Kruskal-Wallis-Test revealed a significant effect between the *none* condition and the *co-driver* condition ($p < 0.05$) as well as between the *none* condition and the *audio* condition ($p < 0.01$). No significant effect could be found between the *audio* condition and the *co-driver* condition. Every participant that aborted the automation, also performed an overtaking action.

Average ratings of simulator sickness over all study groups was 1.90 ($SD = 0.80$) on a 7 point Likert Scale (1 indicating

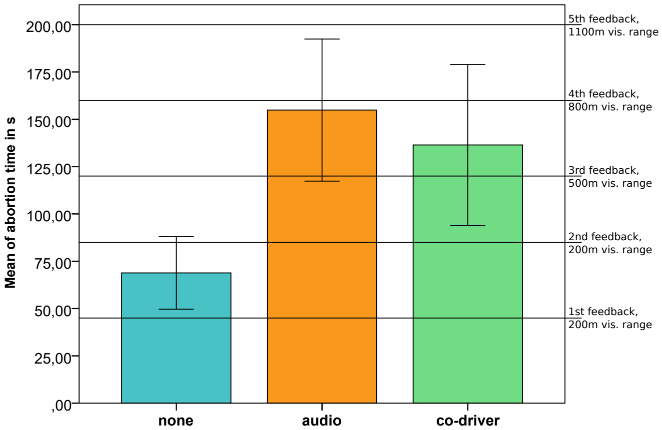

Figure 3. Means of aborting the automation in seconds. The *none* condition with no feedback has the earliest time of abortion. The *audio* condition where audio feedback was given, has the latest time of abortion. In the *co-driver* condition, where the same audio as in the *audio* condition is presented, with an additional virtual co-driver on the front passenger seat, the time of maintaining the automation was between the *none* condition and the *audio* condition. Error bars indicating 95% CI.

no simulator sickness) before driving in the simulation and was rated as 2.11 ($SD = 0.97$) after driving. No significant differences regarding trust and acceptance were found between the three conditions. Immersion was rated as 4.47 ($SD = 0.79$) on average on a seven point Likert Scale (1 indicating no immersion). Trust in the *none* condition was rated as 5.23 ($SD = 0.77$) on a 7 point Likert Scale (1 indicating no trust). Trust in the *audio* condition was rated as 5.31 ($SD = .81$) on average. In the *co-driver* condition, trust was rated as 5.25 ($SD = 0.84$) on average. Acceptance in the *none* condition was rated as 5.58 ($SD = 0.83$) on a 7 point Likert Scale (1 indicating no acceptance), 5.35 ($SD = 1.24$) in the *audio* condition and in the *co-driver* condition on average 5.47 ($SD = 1.25$).

Qualitative feedback regarding the co-driver revealed that the avatar was generally perceived as negative. For most participants it appeared to be uncanny (11) and distracting (7) because head movements were necessary to catch sight of the avatar due to the Oculus Rift's limited horizontal field of view (of 100°in contrast to the natural field of view of about 180°). However, some statements emphasize that avatars may have a positive effect. One participant said that due to the presence of an avatar, he felt more prepared for an interaction between the system and the driver. Another participant reported that he would be more willing to let the avatar finish speaking than in a condition where only voice is applied to provide the feedback.

Discussion

In the *none* condition, the time of aborting the automation and manually overtaking the lead was the shortest. A significant difference between the *none* condition and the *audio* condition as well as between the *none* condition and the *co-driver* condition regarding abortion time was found. H2a and H2b could therefore be confirmed: in the *audio* condition and the *co-driver* condition, the automation is longer maintained than in the *none* condition.

The assumption that an avatar as a social actor has a higher persuasive character than just a voice could not be confirmed. However, the reason for this could be the implementation of the avatar itself. Participants stated that they perceived the character as uncanny. Due to the Oculus Rift's limited field of view, looking at the character was only possible by turning the head, which was perceived as distracting. In real life, the human field of view is about 180 degree. Noticing a passenger on the front passenger seat is easier compared to the VR scene.

Differences in trust and acceptance could also not be found between the three conditions. The assumption that the avatar and system feedback influences trust (H1), could not be confirmed. Consequently, the differences in aborting the automation in the three conditions seem to be not directly related to trust and acceptance.

The low level of simulator sickness indicates that participants did not feel uncomfortable during the study. Generally, a high degree of simulator sickness can lead to the compulsive reaction of finishing the study as soon as possible. This behavior could not be observed. A high level of immersion indicates that participants could identify themselves as being in the scene sitting in the car and performing as good as they could.

The average time difference of 86 seconds between overtaking the car in the *none* condition and the *audio* condition indicates that the delay of abortion between the conditions did not occur because participants were only waiting for the audio to finish before taking action. An audio sample took about 5 to 7 seconds to be played completely.

As pointed out in Section 4.4, a visual range of 635m was required to overtake the car ahead safely. The visual ranges in the study were 200m, 500m, 800m and 1100m. This means that a range of 200m is highly risky, a range of 500m still bears a certain risk and overtaking at a visual range of 800m can be considered as safe driving behavior. The average abortion time in the *none* condition was at 68.8 seconds where visual range accounted for 200m. All but one participant overtook the car at this visual range, one overtook the car at 500m. The average abortion time in the conditions *audio* and *co-driver*, the visual range amounted 500m. This shows that the average of participants in both feedback conditions acted in a safer driving behavior. 35% of participants in the *audio* condition overtook the car not until 800m visual range, likewise 42% in the *co-driver* condition, where overtaking could be considered as safe. In contrast, all of the participants in the *none* condition overtook the car in a potentially dangerous situation. It is thinkable that explicitly telling the participants the actual visual range and that a visual range of about 630m would be necessary to safely overtaking the car, could increase the safe driving behavior. Self-statements of participants after the study suggests that boringness during the drive and the incentive of gaining 3 euros more by driving risky and fast, had a strong influence on most participants to overtake the car ahead early. Although, the influence to overtake the car was high and penalization on potential crashing was low, the differences in aborting the automation between the *none* condition and the feedback conditions are clearly present.

CONCLUSION

In this paper we conducted a user study to investigate system feedback regarding persuasion and safety. System feedback contained the current behavior and an explanation, for example: "The fog is still too dense therefore I cannot overtake the car ahead." Feedback was altered in three different conditions: *none*, where no feedback was given; *audio*, where auditory feedback was generated through text-to-speech; and *co-driver*, where the same *audio* feedback was presented with a virtual co-driver speaking. In the study, participants could decide to continue the automation or abort it and overtake a slow driving lead car in a safety critical driving situation due to a limited visual range. This range increased over time. Study results indicate that with system feedback with and without co-driver participants maintained the automation significantly longer than in the *none* condition. This paper provides evidence that system feedback can persuade drivers in maintaining automation and therefore can lead to an increase of traffic safety. We assume that the use of avatars can lead to an increased level of compliance when sympathy towards the avatar grows. In future work, we expect an increase of persuasion by using a more likable avatar on the dashboard. In this regard, we will pretest avatar sympathy and likability prior to an experimental investigation.

REFERENCES

1. 2016. Definition of PERSUASION. `http://www.merriam-webster.com/dictionary/persuasion`. (2016). Accessed: 2016-05-13.
2. Sean T Doherty, Jean C Andrey, and Carolyn MacGregor. 1998. The situational risks of young drivers: The influence of passengers, time of day and day of week on accident rates. *Accident Analysis & Prevention* 30, 1 (1998), 45–52.
3. Mica R Endsley and Esin O Kiris. 1995. The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 37, 2 (1995), 381–394.
4. Judy J Fleiter, Alexia Lennon, and Barry Watson. 2010. How do other people influence your driving speed? Exploring the who and the how of social influences on speeding from a qualitative perspective. *Transportation research part F: traffic psychology and behaviour* 13, 1 (2010), 49–62.
5. B. J. Fogg. 2002. *Persuasive Technology: Using Computers to Change What We Think and Do* (1 ed.). Science & Technology Books.
6. Google Inc. 2016. Google Self-Driving Car Project. `http://www.google.com/selfdrivingcar`. (2016). Accessed: 2016-04-05.
7. Nicole Gridling, Alexander Meschtscherjakov, and Manfred Tscheligi. 2012. I need help!: exploring collaboration in the car. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work Companion*. ACM, 87–90.
8. Kevin Anthony Hoff and Masooda Bashir. 2015. Trust in automation integrating empirical evidence on factors that influence trust. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57, 3 (2015), 407–434.
9. Jiun-Yin Jian, Ann M Bisantz, and Colin G Drury. 2000. Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics* 4, 1 (2000), 53–71.
10. Nihan Karatas, Soshi Yoshikawa, P Ravindra S De Silva, and Michio Okada. 2015. NAMIDA: Multiparty Conversation Based Driving Agents in Futuristic Vehicle. In *Human-Computer Interaction: Users and Contexts*. Springer, 198–207.
11. A. Kassner and M. Vollrath. 2016. Akzeptanzmessung als Baustein fuer die Entwicklung von Fahrerassistenzsystemen / Acceptance measurement as a component in the development of advanced driver assistance systems. 1960 (2016). `http://trid.trb.org/view.aspx?id=939638`
12. Robert S Kennedy, Norman E Lane, Kevin S Berbaum, and Michael G Lilienthal. 1993. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology* 3, 3 (1993), 203–220.
13. Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry Leifer, and Clifford Nass. 2015. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 9, 4 (2015), 269–275.
14. Manu Kumar and Taemie Kim. 2005. Dynamic Speedometer: Dashboard Redesign to Discourage Drivers from Speeding. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems (CHI EA '05)*. ACM, New York, NY, USA, 1573–1576. DOI: `http://dx.doi.org/10.1145/1056808.1056969`
15. Miltos Kyriakidis, Riender Happee, and JCF De Winter. 2015. Public opinion on automated driving: results of an international questionnaire among 5000 respondents. *Transportation research part F: traffic psychology and behaviour* 32 (2015), 127–140.
16. John Lee and Neville Moray. 1992. Trust, control strategies and allocation of function in human-machine systems. *Ergonomics* 35, 10 (1992), 1243–1270.
17. John D Lee and Katrina A See. 2004. Trust in automation: Designing for appropriate reliance. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 46, 1 (2004), 50–80.
18. Poornima Madhavan and Douglas A Wiegmann. 2007. Similarities and differences between human–human and human–automation trust: an integrative review. *Theoretical Issues in Ergonomics Science* 8, 4 (2007), 277–301.

19. Alexander Meschtscherjakov, David Wilfinger, Thomas Scherndl, and Manfred Tscheligi. 2009. Acceptance of Future Persuasive In-car Interfaces Towards a More Economic Driving Behaviour. In *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '09)*. ACM, New York, NY, USA, 81–88. DOI: http://dx.doi.org/10.1145/1620509.1620526

20. Clifford Nass, Ing-Marie Jonsson, Helen Harris, Ben Reaves, Jack Endo, Scott Brave, and Leila Takayama. 2005. Improving automotive safety by pairing driver emotion and car voice emotion. In *CHI'05 Extended Abstracts on Human Factors in Computing Systems*. ACM, 1973–1976.

21. National Highway Traffic Safety Administration. 2008. National Motor Vehicle Crash Causation Survey: Report to Congress. http://www-nrd.nhtsa.dot.gov/pubs/811059.pdf. (2008). Accessed: 2015-10-23.

22. Michael Nikowitz. 2015. *Fully Autonomous Vehicles: Visions of the future or still reality?* epubli.

23. Rosalind W Picard and Roalind Picard. 1997. *Affective computing*. Vol. 252. MIT press Cambridge.

24. Byron Reeves and Clifford Nass. 1996. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge University Press, New York, NY, USA.

25. John K Rempel, John G Holmes, and Mark P Zanna. 1985. Trust in close relationships. *Journal of personality and social psychology* 49, 1 (1985), 95.

26. Brandon Schoettle and Michael Sivak. 2014. A survey of public opinion about autonomous and self-driving vehicles in the US, the UK, and Australia. (2014).

27. Torsten Schubert, Frank Friedmann, and Holger Regenbrecht. 2001. The experience of presence: Factor analytic insights. *Presence* 10, 3 (2001), 266–281.

28. Bruce G Simons-Morton, Marie Claude Ouimet, Zhiwei Zhang, Sheila E Klauer, Suzanne E Lee, Jing Wang, Rusan Chen, Paul Albert, and Thomas A Dingus. 2011. The effect of passengers and risk-taking friends on risky driving and crashes/near crashes among novice teenagers. *Journal of Adolescent Health* 49, 6 (2011), 587–593.

29. Statista. 2014a. Fully autonomous cars: willingness to buy among U.S. car drivers 2014 | Survey. http://www.statista.com/statistics/428619/us-drivers-who-would-buy-a-fully-autonomous-car/. (2014). Accessed: 2016-05-17.

30. Statista. 2014b. Partially autonomous cars: willingness to buy among U.S. car drivers 2014 | Survey. http://www.statista.com/statistics/428592/us-drivers-who-would-buy-a-partially-autonomous-car/. (2014). Accessed: 2016-05-17.

31. Jinke D Van Der Laan, Adriaan Heino, and Dick De Waard. 1997. A simple procedure for the assessment of acceptance of advanced transport telematics. *Transportation Research Part C: Emerging Technologies* 5, 1 (1997), 1–10.

32. Adam Waytz, Joy Heafner, and Nicholas Epley. 2014. The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology* 52 (2014), 113–117.

33. Kenton Williams and Cynthia Breazeal. 2013. Reducing driver task load and promoting sociability through an affective intelligent driving agent (AIDA). In *Human-Computer Interaction–INTERACT 2013*. Springer, 619–626.

34. Kenton Williams, José Acevedo Flores, and Joshua Peters. 2014. Affective Robot Influence on Driver Adherence to Safety, Cognitive Load Reduction and Sociability. In *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, 1–8.

35. Keith J Williams, Joshua C Peters, and Cynthia L Breazeal. 2013. Towards leveraging the driver's mobile device for an intelligent, sociable in-car robotic assistant. In *Intelligent Vehicles Symposium (IV), 2013 IEEE*. IEEE, 369–376.

36. Jeong-Yean Yang, Yong-Ho Jo, Jae-Chul Kim, and Dong-Soo Kwon. 2013. Affective interaction with a companion robot in an interactive driving assistant system. In *Intelligent Vehicles Symposium (IV), 2013 IEEE*. IEEE, 1392–1397.