Highlights

Effects of Mode Distinction, User Visibility, and Vehicle Appearance on Mode Confusion When Interacting With Highly Automated Vehicles

Mark Colley,Christian Hummler,Enrico Rukzio

- Implementation of a simulation including visualizations.
- Online study with N=59 participants.
- Results indicate that mode distinction and the conspicuous sensor attached to the automated vehicle showed positive effects regarding mode confusion.
- A tintable windshield was negatively assessed.

Effects of Mode Distinction, User Visibility, and Vehicle Appearance on Mode Confusion When Interacting With Highly Automated Vehicles

Mark Colley[a,*], Christian Hummler[a] and Enrico Rukzio[a]

[a]*Institute of Media Informatics, Ulm University, James-Franck-Ring 8, Ulm, 89081, Germany*

ARTICLE INFO

Keywords:
External communication
autonomous vehicles

ABSTRACT

Automated vehicles are expected to communicate with pedestrians at least during the introductory phase, for example, via LED strips, displays, or loudspeakers. While these are added to minimize confusion and increase trust, the human passenger within the vehicle could perform motions that a pedestrian could misinterpret as opposing the vehicle's communication. To evaluate potential solutions to this problem, we conducted an online video-based within-subjects experiment (*N*=59). The solutions under evaluation were mode distinction, vehicle appearance, and the visibility of the passenger via a tintable windshield. Our results show that especially the mode distinction and the conspicuous sensor attached to the automated vehicle showed positive effects. A tintable windshield, however, was negatively assessed. Thus, our work helps to design eHMI concepts to introduce automated vehicles safely by informing about feasible methods to avoid mode confusion.

1. Introduction

In unclear traffic situations, pedestrians and drivers today use explicit (e.g., via hand signals) and implicit (e.g., via deceleration) communication to resolve ambiguities (Lee, Madigan, Giles, Garach-Morcillo, Markkula, Fox, Camara, Rothmueller, Vendelbo-Larsen, Rasmussen et al. (2021)). With automated vehicles (AVs), current research expects the AV to substitute this communication (Colley and Rukzio (2020a)). For this communication to be intuitive and safety-enhancing, numerous modalities and concepts have been proposed (e.g., auditory (Colley, Walch, Gugenheimer, Askari and Rukzio (2020b)), visual (Mahadevan, Somanath and Sharlin (2018)), or tactile concepts (Mahadevan et al. (2018)).

However, most of these works neglect the possible and probable presence of users in AVs. These users will engage in non-driving-related activities (NDRTs) such as sleeping, watching TV, telephoning, and working (Pfleging, Rang and Broy (2016)). With the advance of NDRTs within the AV (games, leisure, work), a wide range of user movement is probable. This passenger movement can easily be misinterpreted. Already it becomes obvious that passengers put their feet out of the window (Edition, 2018, second 3), some played games involving plastic lightsabers (Edition, 2018, second 78), slept (Edition, 2018, second 86), or clapped (Edition, 2018, second 89). For example, a raised hand, while intended for greeting a person in a call or on the street, could be misinterpreted as a warning not to cross the street.

Recent work by Faas et al. (Faas, Stange and Baumann (2021); Faas and Baumann (2021)) shows the need to include status indicators in the external communication, which could reduce or eliminate mode confusion. In the context of AVs, Cummings and Ryan (2014) define mode confusion as a discrepancy in how the human driver believes the vehicle to operate and the actual vehicle behavior. Another relevant factor to consider is that prototypes of AVs and eventually the mass-produced ones do not necessarily resemble current vehicles. For example, the F015 (Benz (2015)) prototype by Mercedes Benz or the BMW Vision M Next (Group (2016)) look very futuristic and, thus, might have lower mode confusion potential due to altered mental models regarding their mode of operation (manual vs. automated) (Dey, Martens, Eggen and Terken (2019)). Also, using additional "conspicuous" sensors such as lidars was already shown to

*Corresponding author

mark.colley@uni-ulm.de (M. Colley); christian.hummler@uni-ulm.de (C. Hummler); enrico.rukzio@uni-ulm.de (E. Rukzio)

https://m-colley.github.io/ (M. Colley)

ORCID(s): 0000-0001-5207-5029 (M. Colley); 0000-0001-5300-9306 (C. Hummler); 0000-0002-4213-2226 (E. Rukzio)

affect the external communication of AVs (Ackermans, Dey, Ruijten, Cuijpers and Pfleging (2020)). Finally, one way to reduce pedestrian-related mode confusion is to block out the passenger, for example, via a tinted windshield.

To evaluate the effect of *mode distinction*, *user visibility*, and *vehicle appearance* and especially their interactions, we conducted a video-based online study with N=59 participants from the USA.

Contribution Statement: Our findings contribute to the development of communication needs for AVs. Insights from our online video-based within-subjects study with N=59 participants show that a tintable windshield creates confusion for pedestrians. While, in line with previous work (Dey et al. (2019)), futuristic-looking vehicles are more considered automated, the willingness to cross was only significantly improved via a distinctive mode or conspicuous sensors attached to the vehicle. We conclude by discussing the relevance for eHMI design.

2. Related Work

This work builds on research in the area of external communication of AVs. As mode confusion is evaluated, we also introduce the terminology and potential measurements and countermeasures.

2.1. External Communication of Automated Vehicles

In manual traffic, traffic-related problems can be overcome via gestures and eye-contact (Rasouli, Kotseruba and Tsotsos (2017)). External communication of AVs, also called external Human-Machine-Interfaces (eHMIs), is a proposed solution to enable such communication between a (potentially) driverless AV and vulnerable road users such as pedestrians or cyclists. While the usage of this necessary explicit communication could be scarce (Lee et al. (2021)), eHMIs could be required in various scenarios.

eHMIs can be classified based on their modality, message type, and communication location (Colley and Rukzio (2020b,a)). Colley and Rukzio (2020a) distinguished eight classes of message type, which refers to the information conveyed by the AV: Instruction, Command, Advisory, Answer, Historical, Predictive, Question, and Affective. The *communication location* defines where the communication occurs: on the vehicle, the personal device, or the infrastructure (Colley and Rukzio (2020a)). Additionally, Situation parameters such as communication relationship (one-to-one, one-to-many, many-to-one, many-to-many), acoustic noise, or communication partner (e.g., pedestrian or cyclist) have to be considered (Colley and Rukzio (2020a)). The work on eHMI focused on children (Deb, Carruth, Fuad, Stanley and Frey (2020); Charisi, Habibovic, Andersson, Li and Evers (2017)), people with vision impairments (Colley et al. (2020b); Colley, Walch, Gugenheimer and Rukzio (2019a)), mobility impairments (Asha, Smith, Freeman, Crump, Somanath, Oehlberg and Sharlin (2021)), cognitive impairments (Haimerl, Colley and Riener (2022)), general pedestrians (Ackermans et al. (2020); Dey, Martens, Wang, Ros and Terken (2018); Löcken, Golling and Riener (2019); Colley, Bajrovic and Rukzio (2022a)), and bicyclists (Hou, Mahadevan, Somanath, Sharlin and Oehlberg (2020)). Additionally, the employed technology was also altered, for example, recent work proposed Augmented Reality glasses-based eHMI approaches (Tabone, Lee, Merat, Happee and de Winter (2021). However, this would require participants to wear such devices and can, in our opinion, not be seen as a requirement but can only be used supplementary. Furthermore, eHMIs were proposed for other use cases than street crossing. For example, they were proposed to aid automated delivery trucks (Colley, Mytilineos, Walch, Gugenheimer and Rukzio (2022b)), to warn distracted pedestrians (Colley, Li and Rukzio (2021c)), and to personalize an AV (Colley, Lanzer, Belz, Walch and Rukzio (2021b)).

Recently, the effects of mode distinction and, as a related result, mode confusion were evaluated. Colley, Wankmüller, Mend, Väth, Rukzio and Gugenheimer (2022c) evaluated which position of the human passenger (driver vs. passenger position), which information type (intention vs. command), and which activity of the passenger (none vs. waving vs. gesturing "stop") lead to mode confusion. Finally, Colley et al. (2022c) propose solutions on the three levels system, user and pedestrian, and the societal level. On the system level, clearly highlighting the status, the vehicle appearance, or the blocking out of the user was proposed.

Dey et al. (2019) investigated the effect of vehicle appearance on the willingness to cross. For this, they used a Renault Twizy. They found that the futuristic appearance did not heighten willingness to cross but actually reduced it. The authors hypothesize that the appearance "interfere[s] with the cognitive models of pedestrians and make them more hesitant in a road-crossing context" (Dey et al., 2019, p. 203).

Faas, Mathis and Baumann (2020b) already stated that "status information on the vehicle's driving mode can be seen as the most basic information" (Faas et al., 2020b, p. 183). While they did not include a condition solely with intent or awareness communication, the answers provided by their participants in the post-experimental interview showed that nearly 90% want to receive information about the AV's automated status. Additionally, Faas and Baumann (2021) found that a status indicator as well as a (simulated) sound have equal effects on subjective measures such as perceived safety and are, therefore, equally important. Finally, Faas et al. (2021) evaluated the status indicator alongside different driver's states (attentive, tinted windshield, distracted). They found no significant effects if the status indicator was present. Regarding crossing onset times, the authors found no significant differences. Faas et al. (Faas and Baumann (2021); Faas et al. (2021)) show promising results for the necessity to include a status indicator. Regarding the color for such a status indicator, turquoise seems to be a sensible color as it is highly visible and does not carry traffic-related meaning until now (Werner (2018)). Despite this work, the role of the tintable windshield and especially the appearance of the vehicle was not taken into account in the work by Faas et al.

The final communication implementation in previous work varies strongly. For example, LED strips (Florentine, Ang, Pendleton, Andersen and Ang (2016); Lundgren, Habibovic, Andersson, Lagström, Nilsson, Sirkka, Fagerlönn, Fredriksson, Edgren, Krupenia and Saluäär (2017)), displays (Florentine et al. (2016)) or projections (Ackermann, Beggiato, Schubert and Krems (2019)) have been used. Other concepts include the smartphone of the pedestrian (Mahadevan et al. (2018); Holländer, Krüger and Butz (2020)) or even a physical waving hand mounted on top of the vehicle (Mahadevan et al. (2018)). Displays are capable to display various messages and include, for example, anthropomorphic (Semcon (2016)) or text (Colley, Mytilineos, Walch, Gugenheimer and Rukzio (2020a); Colley, Belz and Rukzio (2021a)). Previous work found that text was least ambiguous (Chang, Toda, Igarashi, Miyata and Kobayashi (2018); Deb et al. (2020)). Therefore, we employed text to avoid potential confusion based on the eHMI concept. We believe that with more ambiguous eHMI concepts, mode confusion will increase.

2.2. Mode Confusion

Mode confusion was investigated for aircrafts (Spencer Jr (2000); Degani, Shafto and Kirlik (1996)), flight guidance systems (Joshi, Miller and Heimdahl (2003)), service robots (Lankenau (2001)), and driver interfaces (Lee, Ahn and Yang (2014)).

Degani and Kirlik report mode-switches, that is switching the mode of operation from automated to manual flying or taking over control from an AV during an automated journey, in aircrafts to be intuitive when being user-visible (e.g., via a display) (Degani and Kirlik (1995)). These automation modes and their switching is comparable to the six vehicle automation levels (Taxonomy (2014)). Degani et al. (1996) distinguish two mode confusion categories: (1) misidentifying automation behavior, that is the automation operates differently than the user expects, and (2) wrongly assuming an automation mode than is actually active. The second is in line with Norman's definition (Norman (1983)). Leveson, Pinnel, Sandys, Koga and Reese (1997) identify six problems with mode confusion: Interface Interpretation Error, Inconsistent Behavior, Indirect Mode Change, Operator Authority Limits, Unintended Side Effects, and Lack of Appropriate Feedback.

Kurpiers, Biebl, Mejia Hernandez and Raisch (2020) distinguish measuring subjective and objective mode confusion. Subjective measures provide quick and explicit information. However, they inherently include, for example, personal bias. Additionally, system misunderstandings can not be directly unveiled as they are per se, not user-visible. Objective measures include gaze behavior, facial expression, behavior patterns, and engagement in non-driving task-related tasks. The foremost interest in mode awareness is, however, in the following behavior alteration (Kurpiers et al. (2020).

Mode confusion measurement is difficult (Kurpiers et al. (2020)) and, therefore, interviews are often used (Johnson and Pritchett (1995); Kurpiers et al. (2020)). Other measurements include time to recognize a problematic mode (Johnson and Pritchett (1995)) or having participants determine the system's current mode (Lee et al. (2014)).

Cognitive load (or lockup) was hypothesized to be associated with mode error by Spencer Jr (2000). As divergent mode capability assumptions have to be resolved, cognitive resources are necessary. Thus, we assume cognitive load to increase with mode confusion. Cunningham and Regan (2015) focus on AV driver's overtrust. However, they determine trust to be contributing to mode confusion.

With regard to AVs, the passenger within the AV is the main user. However, when including bystanders, as done in the field of eHMIs, with pedestrians or cyclists communicating with the AV, the pedestrians and other vulnerable road users (Holländer, Colley, Rukzio and Butz (2021)) become exposed to potential mode confusion issues when the vehicle state is unclear. For the pedestrian, detailed knowledge about the currently active automation level is irrelevant; however, the binary information *vehicle drives automated* vs. *vehicle is driven manually* (see wrong automation mode assumption (Degani et al. (1996))) is highly important to avoid that crossing decisions are based on false assumptions.

3. Experiment

To evaluate the mode confusion and potential countermeasures, we designed and conducted an online video-based experiment.

3.1. Study Design and Hypotheses

This study investigates the influence of *vehicle appearance*, *mode distinctiveness*, and *passenger visibility* on pedestrian behavior. These are based on work by Faas et al. (2021) that already included *mode distinctiveness* and *passenger visibility* as well as work by Dey et al. (2019) and Ackermans et al. (2020) who evaluated the *vehicle appearance*. With current prototypes having radically different appearances (e.g., Zoox, Mercedes Benz AVTR), we argue that evaluating the effect of *vehicle appearance* is crucial for the studied challenge of mode confusion.
However, Faas et al. were, due to their setup, unable to change the *vehicle appearance*, and Dey et al. as well as Ackermans et al. did not include the mode distinctiveness or passenger visibility. Additionally, the most important distinction to previous work is the focus on mode confusion, which was not directly measured. Therefore, this work provides a holistic approach to these three factors and their interactions.

The research question was:

> *What impact do the variables Vehicle Appearance, Mode Distinctiveness, and Passenger Visibility have on (1) cognitive load, (2) trust, (3) attribution of control, (4) perceived safety, and (5) willingness to cross?*

To answer this research question, we presented participants with a video-based online study in which they encountered different AV appearances with an eHMI. These AVs, depending on the condition, showed their automated mode via a turquoise LED at the top of the windshield (see Figure 1a) and were able to block out the passenger via opaque windshields, a technology currently under development (LUMINQ (2020)) (see Figure 1a). The appearances of the vehicles were conventional (modeled after a Passat B6-2006 Variant), the same Passat but with a conspicuous sensor (see (Ackermans et al. (2020))), and a futuristic AV resembling the F 015 concept by Mercedes Benz as a surrogate (Benz (2015)). The passenger wore glasses resembling a potential future Augmented Reality headset. To create confusion, the passenger performed a mid-air selection gesture, which could be mistaken for a "Stop! Watch out" gesture toward the pedestrian (see Figure 1b and Figure 1c). In all conditions, the eHMI displayed the word "STOPPING".

Therefore, we designed our experiment as a 3 (vehicle appearances) × 2 (mode distinctiveness) × 2 (passenger visibility) within-subjects study.

Our hypotheses (H) were:

- H_1: Pedestrians' mode confusion will be reduced with a person not visible, an LED showing the mode distinctively, or a vehicle appearance indicating the automated nature of the vehicle (see Faas et al. (2021) and Faas and Baumann (2021)).
- H_2: Pedestrians' mode confusion will be reduced with combinations of the person not visible, an LED showing the mode distinctively, or a vehicle appearance indicating the automated nature of the vehicle.
- H_3: Pedestrians' willingness to cross before the AV has stopped will be higher with the better visible changes to the AV's appearance.

The hypotheses were only formulated for the mode confusion and the willingness to cross. For the other dependent variables, the study was exploratory.

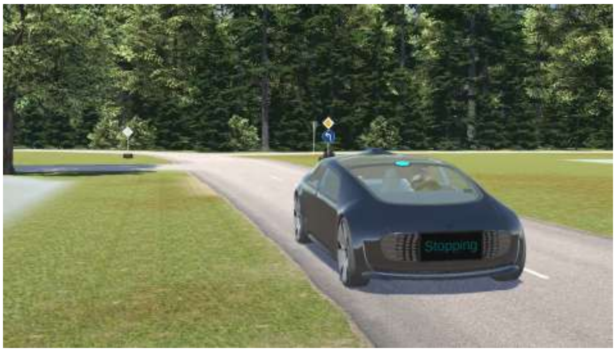
3.2. Materials

We modeled a typical scenario in Unity version 2020.3.1f1 (Unity Technologies (2020)) to enable us to evaluate currently unavailable technology (tintable windows and futuristic vehicle appearances). Before our experiment, we questioned seven participants in informal interviews regarding the external validity (i.e., how real the videos looked). Participants were asked to indicate similarity for the aspects *Colouring + perspective*, *Behaviour of the vehicle and passenger*, and *Own (hypothetical) behavior in the situations*. Based on their responses, we adjusted our simulation. We repeated this twice until all participants were satisfied. The videos took 13s each and depicted the AV driving around a curve and then slowing down linearly. The futuristic vehicle emitted a sound closely resembling the BMW i4 sound by Hans Zimmer[1].

[1] https://www.youtube.com/watch?v=JpeukPXMWzA; Accessed: 21.02.2022

(a) Not visible - distinctive - conventional.

(b) Visible - not distinctive - conspicuous.



(c) Visible - distinctive - futuristic.

Figure 1: Three screenshots of various conditions we used in the online study.

3.3. Measurements

We used the mental workload subscale of the raw NASA-TLX (Hart and Staveland (1988)) on a 20-point scale ("How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?"; 1=Very Low to 20=Very High). The subscales *Predictability/Understandability* (*Understanding* from here) and *Trust* of the *Trust in Automation* questionnaire by Körber (2019) were used. Understanding is measured using agreement on four statements ("The system state was always clear to me.", "I was able to understand why things happened."; two inverse: "The system reacts unpredictably.", "It's difficult to identify what the system will do next.") using 5-point Likert scales (1=Strongly disagree to 5=Strongly agree). Trust is measured via agreement on the same 5-point Likert scale on two statements ("I trust the system." and "I can rely on the system."). Also, participants rated the perceived safety using four 7-point semantic differentials from -3 (anxious/agitated/unsafe/timid) to +3 (relaxed/calm/safe/confident) (Faas, Kao and Baumann (2020a)).

On 7-point Likert scales, participants rated whom they believed to be in control ("Control"; *1=Person* to *7=Vehicle*) and whom their crossing decision was based on ("Decision Factor"; *1=Person* to *7=Display*). Agreement to the statement *While the vehicle was communicating, the behavior of the human passenger confused me; "Confusion"* was rated as well as whether the vehicle's intentions and actions ("vehicle mode clarity") and of the human user ("passenger mode clarity") were clear (each one question; *1=Totally Disagree* to *7=Totally Agree*). We additionally showed participants three screenshots of the video per condition: (1) at the beginning of the video, (2) when the eHMI became visible, and (3) after the vehicle had stopped (these are shown in Figure 1). For every screenshot, participants rated their willingness to cross (*1=completely unsure* to *7=completely sure*). Driving style was assessed using a single item (*1=completely safe* to *7=completely dangerous*). Participants also indicated whether they would cross the street in this scenario *during* the approach and *after* the vehicle stopped (boolean; yes/no).

After having seen all videos, participants answered openly posed questions regarding feedback and improvement possibilities. Additionally, they ranked the scenarios with regards to willingness to cross and confusion. Immersion, which was used to validate the simulation's appropriateness and the overall video-based approach, was measured using the *Immersion* subscale of the Technology Usage Inventory (Kothgassner, Felnhofer, Hauk, Kastenhofer, Gomm and

Kryspin-Exner (2013)) showing medium immersion (*M*=15.91, *SD*=5.49). Participants also rated the realism of the videos ("I found the videos realistic") on a 7-point Likert scale (*1=Totally Disagree* to *7=Totally Agree*), rating the videos realistic (*M*=4.74, *SD*=1.57).

3.4. Procedure

Each session started with the introduction and the agreement of informed consent. The introduction to the study was as follows (including wording in boldness):

*You will see **twelve** videos of a highly automated vehicle approaching you. Future automated vehicles could have an attached display, for example, on the grille to communicate with pedestrians. The vehicle **steers, brakes, and accelerates** (lateral and longitudinal guidance).*
*Passengers will be able to engage in **every imaginable activity**. You are supposed to imagine standing at the curb and assess whether you would be willing to cross in front of the highly automated vehicle.*

Additionally, the mode LED, the tintable window, and the different vehicle appearances were explained. Participants had to indicate their understanding afterward. Then, the videos were presented in randomized order. The explanations regarding eHMI, autonomous driving mode, mode LED, and vehicle appearance were repeated before every condition. A session lasted ≈ 35min; participants were compensated with 4.70€.

4. Results

4.1. Participants

Before the experiment, we computed the required sample size via an a-priori power analysis using G*Power (Faul, Erdfelder, Buchner and Lang (2009)). To achieve a power of .9, with an alpha level of .05, 41 participants should result in small to medium effect size (0.15 (Funder and Ozer (2019)) in a within-subjects repeated measures ANOVA.

Therefore, we recruited 66 participants via Prolific pre-selected from the USA to avoid confounding variables such as left-hand vs. right-hand traffic or culture (Rasouli and Tsotsos (2019)). Seven data sets were excluded due to failed attention checks. The remaining *N*=59 (27 female, 31 male, 1 non-binary) participants were *M*=35.00 (*SD*=11.82) years old. Participants showed medium *Propensity to Trust* (Körber (2019)) with a mean of *M*=2.81 (*SD*=.63). Participants reported medium to high interest in AVs (*M*=3.88, *SD*=.89), medium belief that AVs will ease their lives (*M*=3.63, *SD*=1.08), and rather believed that AVs would become reality in the next 10 years (*M*=4.22, *SD*=1.05)

4.2. Data Analysis

For the factor analysis in the case of non-parametric data, the non-parametric ANOVA (NPAV) as described by Lüpsen (Lüpsen (2020)) was employed. For post-hoc tests, we used Bonferroni correction. R in version 4.1.2 and RStudio in version 2022.02.0 was employed. All packages were up to date in February 2022. Effect sizes were calculated using Rosenthals's formula (Rosenthal, Cooper and Hedges (1994)) unless stated otherwise. The figures show the reported main or interaction effect in bold lines and include 95% confidence intervals.

4.3. Perceived Safety, Trust, & Cognitive Load

The NPAV found no significant effects on perceived safety. The mean values ranged between *M*=0.99 (Tinted window, not distinctive, futuristic) to *M*=1.64 (Not tinted, distinctive, conspicuous).

The NPAV found a significant main effect of *mode distinctiveness* on Understanding ($F(1, 58) = 10.74$, *p*=0.002). The NPAV also found a significant interaction effect of *vehicle appearance* × *mode distinctiveness* × *tintable windshield* on Understanding ($F(2, 116) = 3.48$, *p*=0.034; see Figure 2). While Understanding was always higher with a distinctive mode, the difference with a clear windshield became smaller for the conspicuous and smallest for the futuristic *vehicle appearance*. With a tinted windshield, subjective Understanding was almost equal for the conventional and conspicuous *vehicle appearance* but lower for the futuristic one when not being distinctive.

The NPAV found a significant main effect of *mode distinctiveness* on trust ($F(1, 58) = 7.80$, *p*=0.007, *r*=-0.101, Z=-2.69). Trust was significantly higher when the mode was shown distinctively via a LED (*M*=3.56, *SD*=1.02) than it not being shown (*M*=3.46, *SD*=1.01).

Regarding mental workload, the NPAV found no significant effects. Nonetheless, a few difference almost reached significance: *vehicle appearance* (*p*=0.06, *vehicle appearance* × *mode distinctiveness* (*p*=0.055; see Figure 3a), and *mode distinctiveness* × *tintable windshield* (*p*=0.053; see Figure 3b).

Required cognitive load was low when the mode was distinctive for all AV appearances. The cognitive load was higher for the conventional and the futuristic appearance compared to the conspicuous appearance.

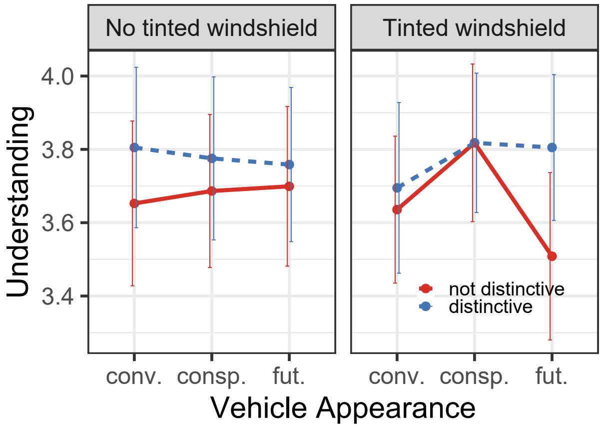
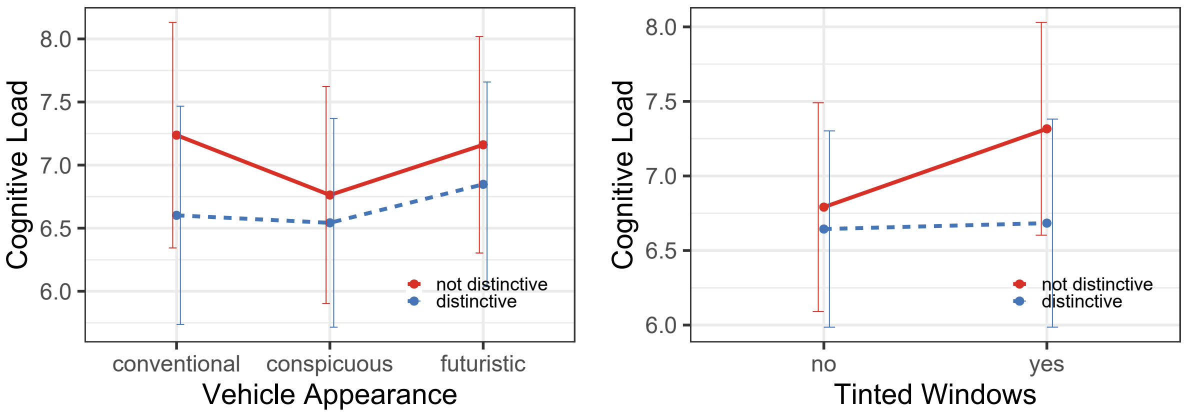


Figure 2: IE *vehicle appearance* × *mode distinctiveness* × *tintable windshield* on Understanding.



(a) IE of *vehicle appearance* × *mode distinctiveness*.

(b) IE of *tintable windshield* × *mode distinctiveness*.

Figure 3: Non-significant IEs on mental workload.

4.4. Willingness to Cross

The NPAV found a significant main effect of *tintable windshield* on willingness to cross in the first screenshot ($F(1, 58)$ = 8.38, p=0.005, r=-0.105, Z=-2.79). Willingness to cross in the first screenshot was higher with a distinctive mode (M=3.81, SD=2.30) than without (M=3.65, SD=2.23).

For the willingness to cross in the second screenshot, the NPAV found a significant effect of *tintable windshield* ($F(1, 58)$ = 8.48, p=0.005, r=-0.105, Z=-2.80). Willingness was higher when the window was not tinted (M=2.96, SD=1.89) than with it being tinted (M=2.76, SD=1.96).

The NPAV found no significant differences in the willingness to cross for the third screenshot.

4.5. Mode Confusion

The NPVA found no significant effect on confusion (measured via agreement to the statement: “While the vehicle was communicating, the behavior of the human passenger confused me”).

The NPAV found a significant main effect of *mode distinctiveness* on vehicle mode clarity ($F(1, 58)$ = 4.50, p=0.038, r=-0.078, Z=-2.07). Subjective vehicle mode was clearer with the LED (M=5.66, SD=1.41) than without (M=5.47, SD=1.44).

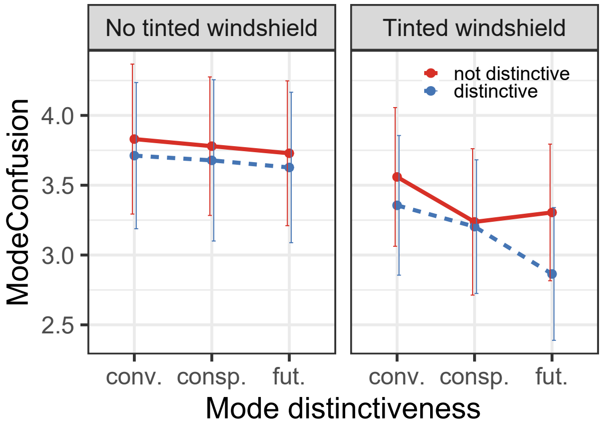
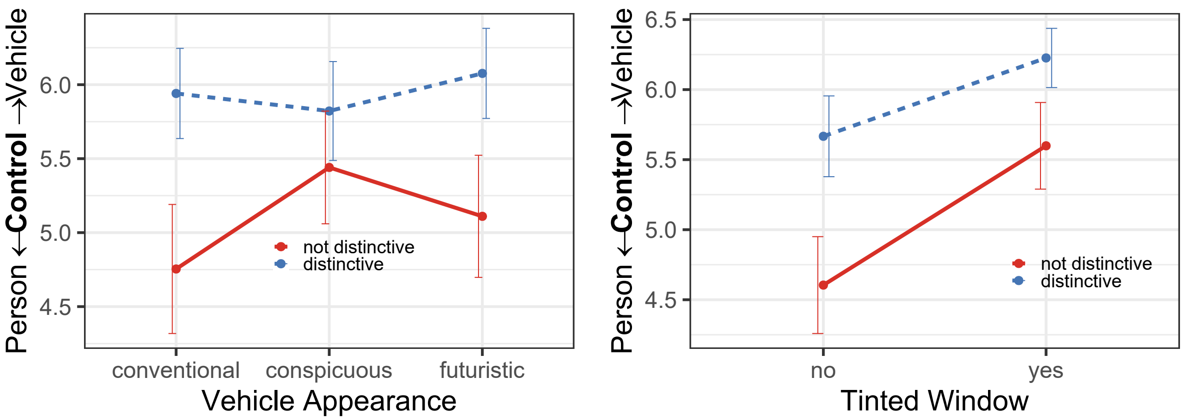


Figure 4: Values of mode confusion (lower is better).

For the passenger mode clarity, the NPAV found a significant effect of *tintable windshield* ($F(1, 58) = 23.55$, p<0.001, r=-0.166, Z=-4.43). Subjective passenger mode was clearer when the window was not tinted (M=4.33, SD=2.00) than with it being tinted (M=3.24, SD=1.93).

4.6. Attribution of Control, Driving Style, and Decision Factor

The NPAV found a significant main effect of *mode distinctiveness* on control attribution ($F(1, 58) = 14.79$, p<0.001, r=-0.136, Z=-3.61). The NPAV found a significant main effect of *tintable windshield* on control attribution ($F(1, 58) = 24.42$, p<0.001, r=-0.169, Z=-4.50).



(a) IE of *vehicle appearance* × *mode distinctiveness*.

(b) IE of *tintable windshield* × *mode distinctiveness*.

Figure 5: IEs on control attribution.

The NPAV found a significant interaction effect of *vehicle appearance* × *mode distinctiveness* on control attribution ($F(2, 116) = 4.80$, p=0.010; see Figure 5a).

The NPAV found a significant interaction effect of *mode distinctiveness* × *tintable windshield* on control attribution ($F(1, 58) = 4.29$, p=0.043; see Figure 5b).

The NPAV found a highly significant main effect of *tintable windshield* on decision factors to cross ($F(1, 58) = 52.29$, p<0.001, r=-0.229, Z=-6.08). The NPAV found a significant interaction effect of *vehicle appearance* × *mode distinctiveness* on decision factors to cross ($F(2, 116) = 3.83$, p=0.024; see Figure 6). The decision was mostly based on the display when the vehicle's appearance was conventional or futuristic and there was an LED for distinction.

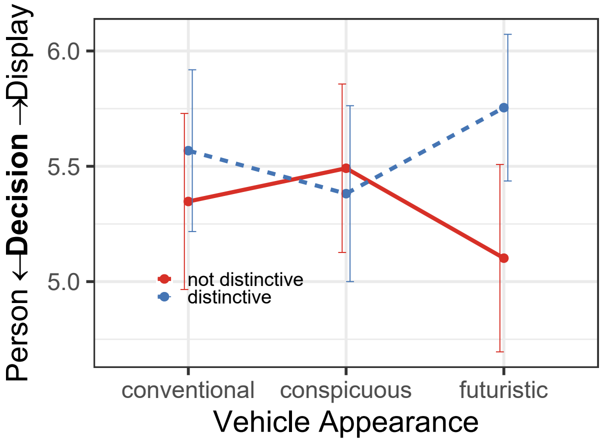


Figure 6: IE of *vehicle appearance* × *mode distinctiveness* on decision factor.

However, when the appearance was conspicuous, participants reported having based their decision on the display more if there was no LED for distinction.

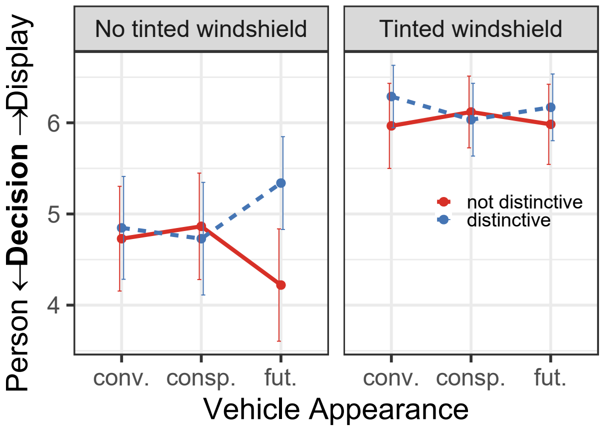


Figure 7: IE of *vehicle appearance* × *mode distinctiveness* × *tintable windshield* on decision factor.

The NPAV also found a significant interaction effect of *vehicle appearance* × *mode distinctiveness* × *tintable windshield* on decision factors to cross ($F(2, 116) = 3.64$, p=0.029; see Figure 7). With a tinted windshield, the decision was based more on the display as the person was then not visible. With the tinted windshield, the decision factor received close values. For the clear windshield, the futuristic vehicle appearance leads to a higher reliance on the display with the distinctive mode than without.

The NPAV found a significant main effect of *vehicle appearance* on driving style ($F(2, 116) = 3.13$, p=0.047, r=-0.074, Z=-1.98). However, a post-hoc test found no significant differences. The NPAV found a significant main effect of *tintable windshield* on driving style ($F(1, 58) = 5.24$, p=0.026, r=-0.084, Z=-2.23). The driving style was rated as safer (M=2.51, SD=1.60) with a clear compared to a tinted windshield (M=2.78, SD=1.65).

4.7. Decision to Cross

As crossing relies on a boolean decision, we asked participants whether they would cross in front of the vehicle **during** the approach and **after** the vehicle stopped. For the analysis, we employed a generalized linear mixed model (GLMM) using ML and BOBYQA optimizer. Due to the within-subjects design, we specified the participant as a

	During Approach		After Stopped	
Variable number of answers	Yes	No	Yes	No
shape				
conventional	70 (29.66%)	166	219 (92.80%)	17
conspicuous	89 (37.71%)	147	223 (94.49%)	13
futuristic	77 (32.63%)	159	220 (93.22%)	16
mode distinctive				
not distinctive	130 (36.72%)	224	338 (95.48%)	16
distinctive	106 (29.94%)	248	324 (91.53%)	30
Passenger Visibility				
Visible	106 (29.94%)	248	325 (91.81%)	29
Not visible	130 (36.72%)	224	337 (95.20%)	17

Table 1
Participants' decision to cross concerning the factors both during the vehicle's approach and after the vehicle stopped. Being Visible means that the windshield was not tinted.

random effect in our model, therefore accounting for individual differences. We also report descriptive values for the crossing decision (see Table 1).

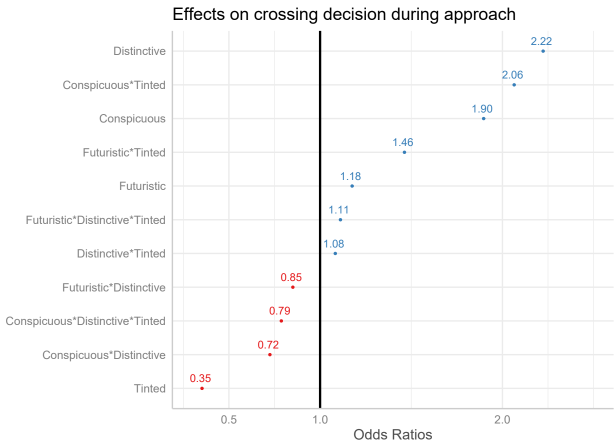


Figure 8: GLMM for the crossing decision during the approach. The odds ratio defines the probability that a participant would have crossed during the approach of the AV. The odds ratio of 2.22 for the *mode distinctiveness* shows that the probability of crossing during the approach was 2.22 times higher compared to the no *mode distinctiveness*. On the other end, the probability of crossing during the approach was only 0.35 times as probable when the windshield was *tinted*.

Regarding Figure 8, the GLMM showed no significant positive or negative effect on the crossing decision *during* the approach. However, the *mode distinctiveness* showed the highest odds ratio (2.22) while *tintable windshield* showed the lowest. The combination of Conspicuous × Tinted also showed a relatively high odds ratio (2.06). Therefore, we assume that the clear and mostly unambiguous nature of the LED and the sensor had the highest impact for the crossing decision during the approach. Not seeing the passenger, however, lead to reduced willingness probably due to confusion.

Regarding Figure 9, the three-way interaction between the vehicle appearance *futuristic*, the *tinted* windshield, and the *distinctive* mode had a significant positive impact on the decision to cross *after* the approach. However, the high number of "No" responses (see Table 1) for the decision *after the AV stopped* and, in general, the three-way interaction, makes interpreting this effect difficult, which is why we refrain from providing an interpretation. The *distinctive* mode alone had a significant positive effect on the decision to cross *after* the AV stopped. This is in line with the findings for the decision *during* the approach.

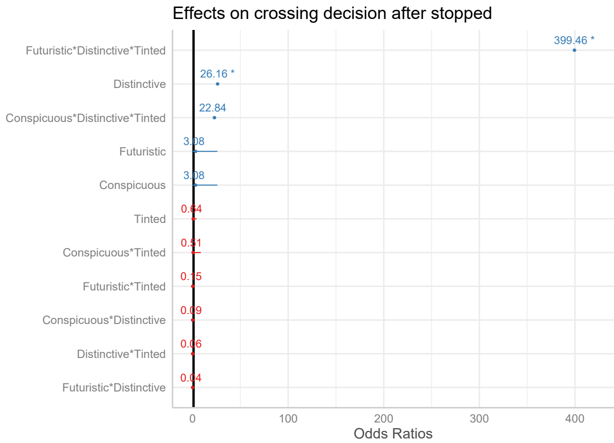


Figure 9: GLMM for the crossing decision after stopping. The odds ratio defines the probability that a participant would have crossed during the approach of the AV. For example, the odds ratio of 26.16 for the *mode distinctiveness* shows that the probability of crossing after the vehicle stopped was 26.16 times higher compared to the no *mode distinctiveness*.

4.8. Reasonability, Necessity, and Open Feedback

After all conditions, participants rated the reasonability and the necessity of the external communication via screen and the *mode distinctiveness* on 7-point Likert scales. Both the reasonability (M=5.88, SD=1.27) and the necessity (M=5.74, SD=1.40) of the external communication were rated high. Participants also rated the reasonability (M=5.95, SD=1.26) and the necessity (M=5.65, SD=1.28) of the *mode distinctiveness* highly.

In the open feedback, three participants stated that the gesture of the passenger confused them (e.g., “The hand signal of the driver, when visible, confused me slightly. It seemed like the driver was signaling, "wait".”). This showed that our manipulation worked. Three participants were undecided whether tinting the window was a good idea (e.g., “ It’s disconcerting not seeing the person, but if I could see they weren’t paying attention that would be disconcerting too.”, “one thing that I could not stand is the windshield being tinted”). The same participant stated their doubts about AVs (and computers in general): “Machines/computers malfunction all the time and it’s just not safe to trust it won’t have an issue.” One participant also stated that the textual message “stopping” was confusing as it was still displayed after having stopped (“Because the car still said stopping, I assumed that the car had not come to a complete stop at any point”). Also, another participant stated that the AV could simply have come a halt earlier to avoid confusion.

5. Evaluation of Hypotheses and Discussion

This work presented possible solutions to the problem of mode confusion in AVs with a human user. Based on previous work (e.g., Faas et al. (2020b)), the solutions on the system level are an easily recognizable *vehicle appearance*, *mode distinctiveness*, and a *tintable windshield*. With the presented data, the hypotheses can be accepted/rejected as follows:

5.1. Evaluation of Hypotheses

- H_1: *Pedestrians’ mode confusion will be reduced with a person not visible, an LED showing the mode distinctively, or a vehicle appearance indicating the automated nature of the vehicle.*
  With *mode distinctiveness* and a *tintable windshield*, participants (correctly) significantly higher attributed control to the AV. The *vehicle appearance* showed, however, no significant main effect. Additionally, we could not show a significant effect of either *vehicle appearance*, *mode distinctiveness*, or *tintable windshield* on the direct measurement on mode confusion. However, we found that a distinctive mode led to higher subjective vehicle mode clarity and a tinted windshield led to lower subjective passenger mode clarity Also, we found various interaction effects, leading us to the conclusion that the factors are, when employed independently, not able to explain mode confusion satisfactorily. Therefore, we partially reject this hypothesis.

- *H_2: Pedestrians' mode confusion will be reduced with combinations of the person not visible, an LED showing the mode distinctively, or a vehicle appearance indicating the automated nature of the vehicle.*
  While values were lowest with the futuristic vehicle appearance, with a tinted windshield, and the mode distinctive (see Figure 4), we could not show any significant IEs on mode confusion or subjective passenger or vehicle mode clarity. Therefore, we reject this hypothesis.

- *H_3: Pedestrians' willingness to cross before the AV has stopped will be higher with the better visible changes to the AV's vehicle appearance.*
  No effects with regards to the *vehicle appearance* were found for the first, second, or third screenshot. Additionally, the GLMM found no significant main effect of vehicle appearance during the approach (see Figure 8) or after the vehicle stopped (see Figure 9).Therefore, we reject the hypothesis.

5.2. Reducing Mode Confusion

With this experiment, we were able to show that crossing decision depends on a variety of factors, including *vehicle appearance*, *mode distinctiveness*, and *tintable windshield*. This is in line with previous work showing that conspicuous sensor (Ackermans et al. (2020)) or a futuristic vehicle appearance (Dey et al. (2019)) altered pedestrians' willingness to cross. However, while Dey et al. (2019) employed a small Renault Twizy and, therefore, discussed whether the size difference to the BMW 3 was relevant to the crossing decision, their data indicate that the size was of little relevance. Therefore, we conclude that in their experiment, the futuristic appearance had the greatest effect with little to no effect of size. In line with the work by Dey et al. (2019), our data showed that control was more attributed to the vehicle with the futuristic appearance, especially with the distinctive mode (see Figure 5b and Figure 5a).

Our data support that the presence of a status indicator (i.e., *mode distinctiveness*) increases Understanding and Trust. This *mode distinctiveness* also increased the willingness to cross and led to a significantly higher attribution of control to the AV. This increased attribution was also present when the windshield was tinted. However, the other effects were not present for the other envisioned factors *vehicle appearance* and *tintable windshield*. While it seems logical to simply exclude the passenger from the communication by blocking the view towards the passenger, this made participants feel uneasy. Also, the futuristic *vehicle appearance* led to a higher cognitive load when no *mode distinctiveness* was present. Our hypothesis H_2 stated that combining the mode confusion countermeasures *vehicle appearance*, *tintable windshield*, and *mode distinctiveness* would lead to the lowest mode confusion. However, our data indicate that the *mode distinctiveness* alone reduces mode confusion best.

5.3. Practical Implications

Our data indicate that for scenarios in which the passenger moves unexpectedly in the AV, it is insufficient to rely on futuristic *vehicle appearance*. Instead, *mode distinctiveness* seems crucial for the pedestrians to understand and trust an AV. As the inclusion of such a LED light is technologically easy, such a LED light should be included in the design of AVs. External communication alone seems to be insufficient for conveying that the vehicle is driving automated. However, our recorded data is based on the precondition that the participants were informed about the LED's meaning. Prior work showed that if this is not the case, people can be confused and infer very different meanings (e.g., warnings or that the vehicle is an emergency vehicle) (Hensch, Neumann, Beggiato, Halama and Krems (2019)). Therefore, standardization and learning are very important aspects that have to be considered for a real-world application. Additionally, from the point of view of the pedestrian, tinted windshields should be avoided.

5.4. Limitations

The sample of N=59 participants was recruited from the USA. As culture plays an essential role in crossing decisions (Rasouli and Tsotsos (2019)) transferability is, therefore, limited. Additionally, the experiment was based on an online video-based study using a Unity simulation and, therefore, lacked several real-world crossing characteristics. Video-based studies lack embodiment, a relevant factor for perceived presence in VR (Rogers, Funke, Frommel, Stamm and Weber (2019)) which was, therefore, proposed to be used in studying AV - pedestrian interaction (Colley, Walch and Rukzio (2019b)). Additionally, the perceived risk of crossing a road is lower in online video-based studies. However, we argue that, based on the reasonable Immersion (M=16.33, SD=5.60) and realism (M=4.89, SD=1.33) scores, the results are valid.

Imagining future AVs was probably not trivial for participants. However, it is likely that some people will be surprised in their first encounters with AVs, therefore, we argue that the presented scenario depicts a possible realistic encounter.

While Kurpiers and Biebl (Kurpiers et al. (2020)) discourage purely subjective ratings for mode confusion measurements in AVs, objective measurements are, without generally available AVs, not yet possible. However, when available, this experiment should be repeated.

While our statistical analysis is valid, some significant results are only marginally significant. Some authors claim that statistical significance should be switched to p<0.005 (Benjamin, Berger, Johannesson, Nosek, Wagenmakers, Berk, Bollen, Brembs, Brown, Camerer et al. (2018)). Nonetheless, the current practice remains that a p<0.05 is seen as statistically significant. Nonetheless, the findings of this work should be validated especially in a real-world scenario.

The alteration of the Passat with conspicuous sensors constitutes a specific design adjustment. Using the F015, however, only resembles **one** specific possibility of a futuristic AV design. Other designs as proposed by Zoox or BMW could elicit different responses. Therefore, the generalization of the findings is limited.

6. Conclusion

Overall, this work evaluated three possible factors that could reduce mode confusion for newly introduced AVs into traffic with external communication: *mode distinctiveness*, *vehicle appearance*, and *tintable windshield*. Mode confusion was operationalized via perceived safety, trust, cognitive load, willingness to cross, confusion statements, and the final decision of whether or not to cross. In an online video-based within-subjects study with N=59 participants, we found that only *mode distinctiveness* increased trust and understanding. Vehicles with conspicuous sensors also received high trust and low cognitive load rating, even without *mode distinctiveness*. This work helps to overcome potential challenges arising from the changing role and actions of people in AVs. Therefore, this work helps introduce AVs safely and adds insights into the required communication design of AVs.

ACKNOWLEDGMENTS

The authors thank all study participants and Sander Ackermans for providing the model for the conspicuous sensor. This work was conducted within the project 'Interaction between automated vehicles and vulnerable road users' (Intuitiver), funded by the Ministry of Science, Research and Arts of the State of Baden-Württemberg.

CRediT authorship contribution statement

Mark Colley: Conceptualization, Methodology, Software, Hardware, Investigation, Formal Analysis, Data Curation, Writing- Original Draft, Writing- Reviewing and Editing, Visualization, Project administration. **Christian Hummler:** Conceptualization, Software, Investigation. **Enrico Rukzio:** Resources, Supervision, Writing- Reviewing and Editing.

References

Ackermann, C., Beggiato, M., Schubert, S., Krems, J.F., 2019. An experimental study to investigate design and assessment criteria: what is important for communication between pedestrians and automated vehicles? Applied ergonomics 75, 272–282.

Ackermans, S., Dey, D., Ruijten, P., Cuijpers, R.H., Pfleging, B., 2020. The effects of explicit intention communication, conspicuous sensors, and pedestrian attitude in interactions with automated vehicles, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. p. 1–14. URL: https://doi.org/10.1145/3313831.3376197, doi:10.1145/3313831.3376197.

Asha, A.Z., Smith, C., Freeman, G., Crump, S., Somanath, S., Oehlberg, L., Sharlin, E., 2021. Co-designing interactions between pedestrians in wheelchairs and autonomous vehicles, in: Designing Interactive Systems Conference 2021, Association for Computing Machinery, New York, NY, USA. p. 339–351. URL: https://doi.org/10.1145/3461778.3462068, doi:10.1145/3461778.3462068.

Benjamin, D.J., Berger, J.O., Johannesson, M., Nosek, B.A., Wagenmakers, E.J., Berk, R., Bollen, K.A., Brembs, B., Brown, L., Camerer, C., et al., 2018. Redefine statistical significance. Nature human behaviour 2, 6–10.

Benz, M., 2015. Der Mercedes-Benz F 015 Luxury in Motion. https://www.mercedes-benz.com/de/innovation/forschungsfahrzeug-f-015-luxury-in-motion/. [Online; accessed: 07-AUGUST-2021].

Chang, C.M., Toda, K., Igarashi, T., Miyata, M., Kobayashi, Y., 2018. A video-based study comparing communication modalities between an autonomous car and a pedestrian, in: Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Association for Computing Machinery, New York, NY, USA. p. 104–109. URL: https://doi.org/10.1145/3239092.3265950, doi:10.1145/3239092.3265950.

Charisi, V., Habibovic, A., Andersson, J., Li, J., Evers, V., 2017. Children's views on identification and intention communication of self-driving vehicles, in: Proceedings of the 2017 Conference on Interaction Design and Children, Association for Computing Machinery, New York, NY, USA. p. 399–404. URL: https://doi.org/10.1145/3078072.3084300, doi:10.1145/3078072.3084300.

Colley, M., Bajrovic, E., Rukzio, E., 2022a. Effects of pedestrian behavior, time pressure, and repeated exposure on crossing decisions in front of automated vehicles equipped with external communication, in: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. pp. 1–11. URL: https://doi.org/10.1145/3491102.3517571, doi:10.1145/3491102.3517571.

Colley, M., Belz, J.H., Rukzio, E., 2021a. Investigating the effects of feedback communication of autonomous vehicles, in: 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Association for Computing Machinery, New York, NY, USA. p. 263–273. URL: https://doi.org/10.1145/3409118.3475133.

Colley, M., Lanzer, M., Belz, J.H., Walch, M., Rukzio, E., 2021b. Evaluating the impact of decals on driver stereotype perception and exploration of personalization of automated vehicles via digital decals, in: 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Association for Computing Machinery, New York, NY, USA. p. 296–306. URL: https://doi.org/10.1145/3409118.3475132.

Colley, M., Li, S., Rukzio, E., 2021c. Increasing pedestrian safety using external communication of autonomous vehicles for signalling hazards, in: Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction, Association for Computing Machinery, New York, NY, USA. pp. 1–10. URL: https://doi.org/10.1145/3447526.3472024.

Colley, M., Mytilineos, S., Walch, M., Gugenheimer, J., Rukzio, E., 2022b. Requirements for the interaction with highly automated construction site delivery trucks. Frontiers in Human Dynamics 4. URL: https://www.frontiersin.org/article/10.3389/fhumd.2022.794890, doi:10.3389/fhumd.2022.794890.

Colley, M., Mytilineos, S.C., Walch, M., Gugenheimer, J., Rukzio, E., 2020a. Evaluating highly automated trucks as signaling lights, in: 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Association for Computing Machinery, New York, NY, USA. p. 111–121. URL: https://doi.org/10.1145/3409120.3410647, doi:10.1145/3409120.3410647.

Colley, M., Rukzio, E., 2020a. A design space for external communication of autonomous vehicles, in: 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Association for Computing Machinery, New York, NY, USA. p. 212–222. URL: https://doi.org/10.1145/3409120.3410646, doi:10.1145/3409120.3410646.

Colley, M., Rukzio, E., 2020b. Towards a design space for external communication of autonomous vehicles, in: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. p. 1–8. URL: https://doi.org/10.1145/3334480.3382844, doi:10.1145/3334480.3382844.

Colley, M., Walch, M., Gugenheimer, J., Askari, A., Rukzio, E., 2020b. Towards inclusive external communication of autonomous vehicles for pedestrians with vision impairments, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. p. 1–14. URL: https://doi.org/10.1145/3313831.3376472, doi:10.1145/3313831.3376472.

Colley, M., Walch, M., Gugenheimer, J., Rukzio, E., 2019a. Including people with impairments from the start: External communication of autonomous vehicles, in: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings, Association for Computing Machinery, New York, NY, USA. p. 307–314. URL: https://doi.org/10.1145/3349263.3351521, doi:10.1145/3349263.3351521.

Colley, M., Walch, M., Rukzio, E., 2019b. For a better (simulated) world: Considerations for vr in external communication research, in: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications: Adjunct Proceedings, Association for Computing Machinery, New York, NY, USA. p. 442–449. URL: https://doi.org/10.1145/3349263.3351523, doi:10.1145/3349263.3351523.

Colley, M., Wankmüller, B., Mend, T., Väth, T., Rukzio, E., Gugenheimer, J., 2022c. User gesticulation inside an automated vehicle with external communication can cause confusion in pedestrians and a lower willingness to cross. Transportation Research Part F: Traffic Psychology and Behaviour 87, 120–137. URL: https://www.sciencedirect.com/science/article/pii/S1369847822000523, doi:https://doi.org/10.1016/j.trf.2022.03.011.

Cummings, M., Ryan, J., 2014. Point of view: who is in charge? the promises and pitfalls of driverless cars.

Cunningham, M., Regan, M.A., 2015. Autonomous vehicles: human factors issues and future research.

Deb, S., Carruth, D.W., Fuad, M., Stanley, L.M., Frey, D., 2020. Comparison of child and adult pedestrian perspectives of external features on autonomous vehicles using virtual reality experiment, in: Stanton, N. (Ed.), Advances in Human Factors of Transportation, Springer International Publishing, Cham. pp. 145–156.

Degani, A., Kirlik, A., 1995. Modes in human-automation interaction: Initial observations about a modeling approach, in: 1995 IEEE International Conference on Systems, Man and Cybernetics. Intelligent Systems for the 21st Century, IEEE. IEEE, New York, NY, USA. pp. 3443–3450.

Degani, A., Shafto, M., Kirlik, A., 1996. Modes in automated cockpits: Problems, data analysis and a modelling framework, in: ISRAEL ANNUAL CONFERENCE ON AEROSPACE SCIENCES, OMANUTH PRESS LTD. OMANUTH PRESS LTD, Tel Aviv, Haifa, Israel. pp. 258–266.

Dey, D., Martens, M., Eggen, B., Terken, J., 2019. Pedestrian road-crossing willingness as a function of vehicle automation, external appearance, and driving behaviour. Transportation research part F: traffic psychology and behaviour 65, 191–205.

Dey, D., Martens, M., Wang, C., Ros, F., Terken, J., 2018. Interface concepts for intent communication from autonomous vehicles to vulnerable road users, in: Adjunct Proceedings of the 10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Association for Computing Machinery, New York, NY, USA. p. 82–86. URL: https://doi.org/10.1145/3239092.3265946, doi:10.1145/3239092.3265946.

Edition, I., 2018. What Happened When Driver Put His Tesla on Auto Pilot? https://www.youtube.com/watch?v=kaAUHpeFj1c&t=53s. [Online; accessed 12-JUNE-2021].

Faas, S.M., Baumann, M., 2021. Pedestrian assessment: Is displaying automated driving mode in self-driving vehicles as relevant as emitting an engine sound in electric vehicles? Applied ergonomics 94, 103425.

Faas, S.M., Kao, A.C., Baumann, M., 2020a. A longitudinal video study on communicating status and intent for self-driving vehicle – pedestrian interaction, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New

York, NY, USA. p. 1–14. URL: https://doi.org/10.1145/3313831.3376484, doi:10.1145/3313831.3376484.

Faas, S.M., Mathis, L.A., Baumann, M., 2020b. External hmi for self-driving vehicles: Which information shall be displayed? Transportation Research Part F: Traffic Psychology and Behaviour 68, 171–186.

Faas, S.M., Stange, V., Baumann, M., 2021. Self-driving vehicles and pedestrian interaction: Does an external human-machine interface mitigate the threat of a tinted windshield or a distracted driver? International Journal of Human–Computer Interaction 37, 1–11. URL: https://doi.org/10.1080/10447318.2021.1886483, doi:10.1080/10447318.2021.1886483, arXiv:https://doi.org/10.1080/10447318.2021.1886483.

Faul, F., Erdfelder, E., Buchner, A., Lang, A.G., 2009. Statistical power analyses using g* power 3.1: Tests for correlation and regression analyses. Behavior research methods 41, 1149–1160.

Florentine, E., Ang, M.A., Pendleton, S.D., Andersen, H., Ang, M.H., 2016. Pedestrian notification methods in autonomous vehicles for multi-class mobility-on-demand service, in: Proceedings of the Fourth International Conference on Human Agent Interaction, Association for Computing Machinery, New York, NY, USA. p. 387–392. URL: https://doi.org/10.1145/2974804.2974833, doi:10.1145/2974804.2974833.

Funder, D.C., Ozer, D.J., 2019. Evaluating effect size in psychological research: Sense and nonsense. Advances in Methods and Practices in Psychological Science 2, 156–168. URL: https://doi.org/10.1177/2515245919847202, doi:10.1177/2515245919847202, arXiv:https://doi.org/10.1177/2515245919847202.

Group, B., 2016. BMW Group - THE NEXT 100 YEARS. https://www.bmwgroup.com/en/company/the-next-100-years.html. [Online; accessed: 07-AUGUST-2021].

Haimerl, M., Colley, M., Riener, A., 2022. Evaluation of common external communication concepts of automated vehicles for people with intellectual disabilities, in: Proc. ACM Hum.-Comput. Interact., Association for Computing Machinery, New York, NY, USA. URL: https://doi.org/10.1145/3546717, doi:10.1145/3546717. accepted.

Hart, S.G., Staveland, L.E., 1988. Development of nasa-tlx (task load index): Results of empirical and theoretical research, in: Advances in psychology. Elsevier, Amsterdam, The Netherlands. volume 52, pp. 139–183.

Hensch, A.C., Neumann, I., Beggiato, M., Halama, J., Krems, J.F., 2019. Effects of a light-based communication approach as an external hmi for automated vehicles–a wizard-of-oz study. Transactions on Transport Sciences 10, 18—32. doi:10.5507/tots.2019.012.

Holländer, K., Colley, M., Rukzio, E., Butz, A., 2021. A taxonomy of vulnerable road users for hci based on a systematic literature review, in: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. pp. 1–13. URL: https://doi.org/10.1145/3411764.3445480.

Holländer, K., Krüger, A., Butz, A., 2020. Save the smombies: App-assisted street crossing, in: 22nd International Conference on Human-Computer Interaction with Mobile Devices and Services, Association for Computing Machinery, New York, NY, USA. pp. 1–11. URL: https://doi.org/10.1145/3379503.3403547, doi:10.1145/3379503.3403547.

Hou, M., Mahadevan, K., Somanath, S., Sharlin, E., Oehlberg, L., 2020. Autonomous vehicle-cyclist interaction: Peril and promise, in: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. p. 1–12. URL: https://doi.org/10.1145/3313831.3376884, doi:10.1145/3313831.3376884.

Johnson, E.N., Pritchett, A.R., 1995. Experimental study of vertical flight path mode awareness. IFAC Proceedings Volumes 28, 153–158.

Joshi, A., Miller, S.P., Heimdahl, M.P., 2003. Mode confusion analysis of a flight guidance system using formal methods, in: Digital Avionics Systems Conference, 2003. DASC'03. The 22nd, IEEE. IEEE, New York, NY, USA. pp. 2–D.

Körber, M., 2019. Theoretical considerations and development of a questionnaire to measure trust in automation, in: Bagnara, S., Tartaglia, R., Albolino, S., Alexander, T., Fujita, Y. (Eds.), Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018), Springer International Publishing, Cham. pp. 13–30.

Kothgassner, O.D., Felnhofer, A., Hauk, N., Kastenhofer, E., Gomm, J., Kryspin-Exner, I., 2013. Technology usage inventory (tui).

Kurpiers, C., Biebl, B., Mejia Hernandez, J., Raisch, F., 2020. Mode awareness and automated driving—what is it and how can it be measured? Information 11, 277.

Lankenau, A., 2001. Avoiding mode confusion in service-robots, in: Integration of Assistive Technology in the Information Age, Proc. of the 7th Int'l Conf. on Rehabilitation Robotics, IOS Press, Amsterdam, The Netherlands. pp. 162–167.

Lee, S.H., Ahn, D.R., Yang, J.H., 2014. Mode confusion in driver interfaces for adaptive cruise control systems, in: 2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC), IEEE. IEEE, New York, NY, USA. pp. 4105–4106.

Lee, Y.M., Madigan, R., Giles, O., Garach-Morcillo, L., Markkula, G., Fox, C., Camara, F., Rothmueller, M., Vendelbo-Larsen, S.A., Rasmussen, P.H., et al., 2021. Road users rarely use explicit communication when interacting in today's traffic: implications for automated vehicles. Cognition, Technology & Work 23, 367–380.

Leveson, N., Pinnel, L.D., Sandys, S.D., Koga, S., Reese, J.D., 1997. Analyzing software specifications for mode confusion potential, in: Proceedings of a workshop on human error and system development, Glasgow Accident Analysis Group. Glasgow Accident Analysis Group, Glasgow, Scotland. pp. 132–146.

Löcken, A., Golling, C., Riener, A., 2019. How should automated vehicles interact with pedestrians? a comparative analysis of interaction concepts in virtual reality, in: Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Association for Computing Machinery, New York, NY, USA. p. 262–274. URL: https://doi.org/10.1145/3342197.3344544, doi:10.1145/3342197.3344544.

LUMINQ, 2020. In-glass displays for improved automotive safety. https://www.lumineq.com/applications/automotive. [Online; accessed: 12-AUGUST-2021].

Lundgren, V.M., Habibovic, A., Andersson, J., Lagström, T., Nilsson, M., Sirkka, A., Fagerlönn, J., Fredriksson, R., Edgren, C., Krupenia, S., Saluäär, D., 2017. Will there be new communication needs when introducing automated vehicles to the urban context?, in: Stanton, N.A., Landry, S., Di Bucchianico, G., Vallicelli, A. (Eds.), Advances in Human Aspects of Transportation, Springer International Publishing, Cham. pp. 485–497.

Lüpsen, H., 2020. R-Funktionen zur Varianzanalyse. http://www.uni-koeln.de/~luepsen/R/. [Online; accessed 25-SEPTEMBER-2020].

Mahadevan, K., Somanath, S., Sharlin, E., 2018. Communicating awareness and intent in autonomous vehicle-pedestrian interaction, in: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, New York, NY, USA. p. 1–12. URL: https://doi.org/10.1145/3173574.3174003, doi:10.1145/3173574.3174003.

Norman, D.A., 1983. Design rules based on analyses of human error. Communications of the ACM 26, 254–258.

Pfleging, B., Rang, M., Broy, N., 2016. Investigating user needs for non-driving-related activities during automated driving, in: Proceedings of the 15th International Conference on Mobile and Ubiquitous Multimedia, Association for Computing Machinery, New York, NY, USA. p. 91–99. URL: https://doi.org/10.1145/3012709.3012735, doi:10.1145/3012709.3012735.

Rasouli, A., Kotseruba, I., Tsotsos, J.K., 2017. Understanding pedestrian behavior in complex traffic scenes. IEEE Transactions on Intelligent Vehicles 3, 61–70.

Rasouli, A., Tsotsos, J.K., 2019. Autonomous vehicles that interact with pedestrians: A survey of theory and practice. IEEE Transactions on Intelligent Transportation Systems 21, 900–918.

Rogers, K., Funke, J., Frommel, J., Stamm, S., Weber, M., 2019. Exploring Interaction Fidelity in Virtual Reality: Object Manipulation and Whole-Body Movements. Association for Computing Machinery, New York, NY, USA. p. 1–14. URL: https://doi.org/10.1145/3290605.3300644.

Rosenthal, R., Cooper, H., Hedges, L., 1994. Parametric measures of effect size. The handbook of research synthesis 621, 231–244.

Semcon, 2016. Self driving car that sees you. https://semcon.com/smilingcar/. Accessed: 2021-11-17.

Spencer Jr, C.F., 2000. Cockpit automation and mode confusion: The use of auditory inputs for error mitigation. Technical Report. AIR COMMAND AND STAFF COLL MAXWELL AFB AL.

Tabone, W., Lee, Y.M., Merat, N., Happee, R., de Winter, J., 2021. Towards future pedestrian-vehicle interactions: Introducing theoretically-supported ar prototypes, in: 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Association for Computing Machinery, New York, NY, USA. p. 209–218. URL: https://doi.org/10.1145/3409118.3475149.

Taxonomy, S., 2014. Definitions for terms related to on-road motor vehicle automated driving systems. Technical Report. Technical report, SAE International.

Unity Technologies, 2020. Unity. https://unity.com/.

Werner, A., 2018. New colours for autonomous driving: An evaluation of chromaticities for the external lighting equipment of autonomous vehicles. URL: https://open-journals.uni-tuebingen.de/ojs/index.php/tct/article/view/692, doi:10.25538/tct.v0i1.692. number: 1.

Mark Colley is a third-year PhD student at Ulm University in Germany. He is interested in the interaction with automation, in particular with automated vehicles. A main focus of his work lays in leveraging automation to support accessibility. Mark holds a Master's Degree in Computer Science from Ulm University. Contact him at mark.colley@uni-ulm.de

Christian Hummler holds a Bachelor's Degree in Media Computer Science from Ulm University. Contact him at christian.hummler@uni-ulm.de

Enrico Rukzio is a Full Professor with the Institute of Media Informatics, Ulm University, Ulm, Germany. His research interests include designing intelligent interactive systems that enable people to be more efficient, satisfied, and expressive in their daily lives. Enrico Rukzio received the Ph.D. degree in computer science from the University of Munich, Munich, Germany. Contact him at enrico.rukzio@uni-ulm.de