



# Article Anthropomorphic Design and Self-Reported Behavioral Trust: The Case of a Virtual Assistant in a Highly Automated Car

Clarisse Lawson-Guidigbe <sup>1,2</sup>, Kahina Amokrane-Ferka <sup>1,\*</sup>, Nicolas Louveton <sup>3</sup>, Benoit Leblanc <sup>2</sup>, Virgil Rousseaux <sup>1</sup> and Jean-Marc André <sup>2</sup>

- <sup>1</sup> IRT SystemX, 91120 Palaiseau, France
- <sup>2</sup> Laboratoire IMS CNRS UMR 5218, Bordeaux INP-ENSC, Université de Bordeaux, 33400 Talence, France
- <sup>3</sup> CeRCA CNRS UMR 7295, Université de Poitiers, Université François-Rabelais de Tours, 86073 Poitiers, France
  - Correspondence: kahina.amokrane-ferka@irt-systemx.fr

Abstract: The latest advances in car automation present new challenges in vehicle-driver interactions. Indeed, acceptance and adoption of high levels of automation (when full control of the driving task is given to the automated system) are conditioned by human factors such as user trust. In this work, we study the impact of anthropomorphic design on user trust in the context of a highly automated car. A virtual assistant was designed using two levels of anthropomorphic design: "voiceonly" and "voice with visual appearance". The visual appearance was a three-dimensional model, integrated as a hologram in the cockpit of a driving simulator. In a driving simulator study, we compared the three interfaces: two versions of the virtual assistant interface and the baseline interface with no anthropomorphic attributes. We measured trust versus perceived anthropomorphism. We also studied the evolution of trust throughout a range of driving scenarios. We finally analyzed participants' reaction time to takeover request events. We found a significant correlation between perceived anthropomorphism and trust. However, the three interfaces tested did not significantly differentiate in terms of perceived anthropomorphism while trust converged over time across all our measurements. Finally, we found that the anthropomorphic assistant positively impacts reaction time for one takeover request scenario. We discuss methodological issues and implication for design and further research.

**Keywords:** human–machine interface; virtual assistant; trust; anthropomorphism; test methods; user studies; evaluation; autonomous car

# 1. Introduction

During the last four decades, the automotive industry has introduced important innovations in the automation of driving tasks. These are often driver assistance systems which, depending on their level, take over some driving tasks or all of them. The Society of Automotive Engineers (SAE) proposes a six-level description of driving automation ranging from fully manual (L0) to fully automated (L5; see Appendix B) driving. The highest levels of automation (L3 or higher), although not yet available to the public, present significant challenges for the acceptance and adoption of this technology. Indeed, despite the benefits of automated vehicles and predictions of the rate at which they will be adopted, barriers remain [1,2]. These include the possible consequences of automation on the driver's mental workload [3] and situational awareness [4,5] but mainly on the lack of user trust [6].

Trust has been widely discussed in the literature and has different definitions depending on the authors. It can be considered as a psychological state [7], a behavioral intention [8], or a behavior [9]. In this paper, we rely on the definition of trust proposed in [10]: "a psychological state of expectation resulting from knowledge and evaluations related to the operator's trust referent in a specific context and guiding his/her decision to use automatic control". Trust in automated systems is measured using different tools across



Citation: Lawson-Guidigbe, C.; Amokrane-Ferka, K.; Louveton, N.; Leblanc, B.; Rousseaux, V.; André, J.-M. Anthropomorphic Design and Self-Reported Behavioral Trust: The Case of a Virtual Assistant in a Highly Automated Car. *Machines* 2023, *11*, 1087. https://doi.org/ 10.3390/machines11121087

Academic Editors: Bo Yang, Zheng Wang and Shuo Cheng

Received: 20 October 2023 Revised: 1 December 2023 Accepted: 7 December 2023 Published: 13 December 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). studies. In [11,12] the authors propose a detailed literature review of trust measurement tools which include self-reported scales, behavioral measures, and physiological measures. The most widely used self-reported scale is from [13]. This scale includes a wide range of factors (dependability, reliability, confidence familiarity, etc.) and has been used for a wide range of automated systems. Other scales are also developed to suit specific systems such as autonomous vehicles [14,15] or robotics [16]. Behavioral measures of trust involve observation and recording of participants' behavior during interaction with the system and interpretation of these behaviors as a trust indicator. For example, a higher trust level in the automated system is equated with an attenuated startle response in a risky situation [14], less monitoring of the system (reliance behavior) [17], or the withdrawing of the participant's own decision to comply with automation decision (compliance behavior) [18]. Similarly, physiological measures such as lower heart rate [14], gaze behavior [19], or electrodermal response [20] are used to indirectly assess participants' trust level.

Trust, and more specifically under-trust, has been identified as one of the most important challenges to the adoption of automated vehicles [6]. Under-trust presents a risk of rejection of driving automation and, therefore, threatens its large-scale adoption and the possibility of societal benefit from its advantages. Under-trust is the opposite of over-trust, which itself presents other challenges, e.g., overdependence on automation and safety risks in situations where the automated system does not function as it should [21].

According to some authors, well-calibrated trust would facilitate the acceptance of automated systems [22] and reduce the risks of misuse and non-use (disuse) [23]. For automated vehicles, some studies show that the solution to under-trust is integrating anthropomorphic design into the vehicle's human–machine interface [24,25].

Anthropomorphism represents, according to various authors, the tendency for humans to attribute human abilities and characteristics to machines or inanimate objects, such as personality, feelings, rational thinking [26,27], or intentions [28]. Consequently, machines are treated as entities (Media Equation Theory) capable of engaging in social interactions [26]. Some authors [29] explain that assigning human characteristics to a machine would make it more familiar, more explainable, and more predictable. Anthropomorphic design, whose goal is to elicit an anthropomorphic perception in the user, simulates life in inanimate objects through design [30,31]. In robotics, three components of anthropomorphic design are identified in [32]: (1) The form of the robot; (2) Its behavior; (3) Its interaction and communication with humans. Anthropomorphic design in automated vehicle interfaces [33,34] makes use of different social stimuli such as visual appearance, voice and natural language, and non-verbal behaviors.

This work focused on a specific type of virtual assistant interface. Our goal was to study the influence of anthropomorphic attributes (voice, natural language, and visual appearance), alone or combined, on the perception of anthropomorphism and trust in the automated system. This led us to the following research questions:

(Q1): Does increasing anthropomorphic attributes lead to an increase in the perception of anthropomorphism?

(Q2): Does increasing anthropomorphic attributes lead to an increase in the level of trust in the automated system?

(Q3): Does the perception of anthropomorphism correlate with trust?

(Q4): Does increasing anthropomorphic attributes lead to better driving performance?

To answer these questions, three user interfaces with different levels of anthropomorphism were designed and tested: baseline, vocal (natural language), and visual assistant (natural language combined with visual appearance).

## 2. Materials and Methods

2.1. Driving Simulation and Automated Driving System (Environment)

The study was conducted in a fix-base driving simulator housing a cockpit. This cockpit was based on a Renault Espace designed to represent a vehicle equipped with two automation levels: manual level (L0) and autonomous level (L3). The autonomous level

can be activated on a highway or expressway and can replace the driver to speeds up to 110 km/h. This study more specifically focused on handover (transition from L0 to L3) and handback (transition from L3 to L0) situations. Different driving scenarios were used in the tests (Table 1). The driving simulation was generated using AVSimulation SCANeR Studio 1.9 software and projected on a 180° hemispheric screen to provide a driving experience as immersive and as ecologically valid (i.e., generalizable toward real-world application) as possible (Figure 1).

Scenario	Description	Related Research Question and Aims
	Training phase	
Training scenario 1	This scenario allows the participant to discover the different user interfaces of the cockpit and the activation/deactivation procedures for the autonomous mode.	
Training scenario 2	This scenario allows the participant to discover the hand back procedure (takeover requested by the system). In this situation, the driver has 60 s to take over the driving.	
	Experimental phase	
Activation Scenario	This scenario is the first of the experimental phase. First, the driver must merge onto a highway in manual mode before activating the autonomous mode in light traffic. It aims to evaluate how increasing anthropomorphic attributes in the interface increases participants' performance, specifically their ability to activate the autonomous mode.	Q1, Q2, Q3, and Q4.
Takeover request in 60 s (TOR 60)	The takeover request in 60 s (TOR 60) scenario starts with 6 min of automated driving on the same road as in the activation scenario. It ends with a non-urgent takeover request during which the participant has 60 s to regain control of the vehicle. This request is triggered by the system and due to the presence of a highway exit on the path followed by the vehicle.	Q1, Q2, Q3, and Q4. It aims to evaluate how increasing anthropomorphic attributes in the interface impact participants' performance, specifically their ability to handle a non-urgent takeover request, and also the variation in their trust and perception of anthropomorphism.
Takeover request in 10 s (TOR 10)	The takeover request in 10 s (TOR 10) scenario starts with 4 min of automated driving on a straight highway in moderately dense traffic. It ends with an urgent takeover request during which the participant has only 10 s to regain control of the vehicle in a complex situation (dense traffic, with a vehicle stopped in the participant's vehicle lane vehicle lane and another in the left blind spot). This request is triggered by the system and due to a sensor failure.	Q1, Q2, Q3, and Q4. It aims to evaluate how increasing anthropomorphic attributes in the interface impact participants' performance, specifically their ability to handle an urgent takeover request, and also the variation in their trust and perception of anthropomorphism.

Table 1. The different scenarios.

The simulator is also used to test multimodal human–machine interface (HMI) configurations. To do so, the cockpit is equipped with screens (cluster, central display, head-up display, steering wheel display, and mirrors), LEDs (side windows, windshield, and steering wheel), speakers (cockpit and headrest), and seat actuators. Haptic feedback in the steering wheel can also be controlled using a Sensodrive motor.



**Figure 1.** (**a**) Simulation environment and cockpit, (**b**) view of the interior of the cockpit.

# 2.2. Experimantal Factors

Three interfaces were compared: baseline (Figure 2); vocal assistant, the interface integrating the voice of the assistant (Figure 3); and visual assistant, an interface integrating a combination of a visual representation and the voice of the assistant (Figure 4). Each group of participants tested only one of the three interfaces.



Figure 2. Baseline interface.



Figure 3. Baseline interface enriched with vocal assistant.



Figure 4. Baseline interface base and vocal assistant enriched with the visual representation.

# 2.2.1. Baseline Interface

This is a multimodal interface (Figure 2) that integrates the following:

- 1. Visual interface including, on the dashboard, information on speed, driving mode availability, and time available in automated mode. On the steering wheel, a small integrated screen allows the driver to change driving modes (automated to/from manual) and displays the status of the autonomous mode (available, activated, and deactivated). The head-up display (HUD) replicates most of the dashboard's information. In the middle of the windshield, an icon replicates the status of the autonomous mode and time remaining before having to take over the driving.
- 2. Sound interface comprising alert tones with different tonality and rhythms played locally in the driver's headrest.
- 3. Haptic messages are created through the integration of actuators in the backrest and seat of the driver.

# 2.2.2. The "Vocal Assistant" Interface

The vocal assistant corresponds to the baseline condition enriched by vocal messages emitted by the assistant. These messages are complete sentences that include actions that the driver must perform and the associated explanations (see examples in Table 2). They are pre-recorded using "Any Text to Voice" software (version 3.6). The vocal messages are spatialized, emitted from a tablet in the middle of the dashboard which integrates the assistant. In this condition, the assistant uses the same iconic visual representation in the windshield as in the "baseline" interface (Figure 2).

Table 2. Example of vocal assistant messages.

Situation	Message Example	Message Type
AD activated	You can regain control at any time. This is only possible if your hands are on the steering wheel and your eyes are on the road. I wish you an excellent journey.	Information
TOR 60 at 60 s before takeover	You must take over control to take the next exit. I won't be able to drive in 60 s.	Explanation
TOR 60 at 10 s before takeover	Take over control.	Explanation
TOR 10	Sensor failure detected; vehicle stops; take over control!	Explanation

## 2.2.3. The "Visual Assistant" Interface

The visual assistant interface combines baseline and vocal assistant interfaces with a humanoid visual representation [25]. This representation is projected in three dimensions using a tablet and a pyramidal structure that generates a hologram that appears in the middle of the dashboard when autonomous mode is activated (Figure 4).

## 2.3. Participants

Thirty-six people participated in the experiment. Following a between-subjects experimental design, three groups of twelve were formed and each group tested only one of the three interface conditions. Participants were selected based on the following criteria: minimum 3 years of driving experience, regular driving activity (at least 3 times a week) on several types of infrastructures (city, highway, and expressway), and minimal knowledge on driving assistance systems. In addition, participants had to answer a question regarding their propensity to use automation technology. Participants who showed no interest in the technology were rejected. Other exclusion criteria, such as motion sickness or simulator sickness as well as visual and/or hearing problems were applied. In the three groups, a gender and age balance were respected.

### 2.4. Procedure

Individual test sessions lasted 2 to 3 h and were divided into five steps (see Figure 5). The first step or "welcome session" of participants contained a detailed description of the goal of the experiment and the data that were collected. The second step consisted of training participants to use the driving simulator through a computer tutorial whose aim was to explain how the automated driving mode works (including activation and deactivation procedures), the vehicle user interfaces, and the operating limits of the autonomous mode. The tutorial was followed by 2 training scenarios (as shown in Table 1) and finally a trust survey measuring global and multidimensional trust [10]. The third step was the experimental phase during which participants had to complete four driving scenarios (see Table 2). After each scenario, participants completed an elicitation interview [35] and a trust survey. The trust survey at this stage was measuring global and multidimensional trust (adapted from [36]; see (Table 3) and Appendix A. In the fourth step, at the end of the experiment, participants completed the perceived anthropomorphism survey [14] and the global trust survey [37]. The fifth and final step was an informal discussion to collect participants' feelings and thank them.



Figure 5. Test session procedure.

# 2.5. Tasks

To avoid bias as much as possible, the following instructions were read to the participant before each driving scenario of the experimental phase:

"The driving scenario starts at on a highway ramp. I ask you to start the vehicle and merge onto the highway. Once on the highway, activate the autonomous mode as soon as it is proposed. When you have completed the activation, I will give you a smartphone that you will use to play the 2048 game. If the takeover alert is triggered at any time during the scenario, you must stop the game and regain driving control as soon as possible".

Trust Scale Used before Experiment [10] (Initial and Post-Training)	Trust Scale Used During Experiment [36] (after Each Driving Scenario)	Trust Scale Used at the End of Experiment [37]
<ul> <li>1—Do you think you know the automated car?</li> <li>2—How much trust would you have in the automated car?</li> <li>3—Do you think the automated car is useful?</li> <li>4—Do you think the automated car is useful?</li> <li>4—Do you think the automated car is necessary?</li> <li>5—Do you think that the automated car would interfere with your usual driving?</li> <li>6—In everyday life, you tend to take risks</li> <li>7—You tend to trust people</li> <li>8—You believe that trusting someone means being able to trust them with something to do</li> <li>9—You believe that trusting someone means being able to trust them with something to do</li> <li>9—You believe that it is necessary to know a person well to trust him/her</li> <li>10—You think that the decision to trust someone depends on how you interact with them</li> <li>11—You are generally suspicious of new technologies (cell phones, computers, internet, microwave ovens, etc.)</li> <li>12—You think that new technologies are interesting</li> <li>13—You think that new technologies are dangerous</li> <li>What risks would you associate with using an automated car?</li> <li>14—The risk of an accident with the car ahead</li> <li>17—The risk of an accident with the car ahead</li> <li>17—The risk of an accident with the car ahead</li> <li>17—The risk of being dependent on the system</li> <li>19—The risk of losing the pleasure of driving</li> <li>What benefits would you get from using an automated car?</li> <li>20—Less stressful driving</li> <li>21—Lightening the driving task (the automated car would allow me to perform a secondary activity such as reading or using my phone)</li> <li>22—Easier driving task</li> <li>23—Improved driving comfort</li> <li>24—Safer driving</li> </ul>	<ul> <li>1—I trust the automated car in this situation</li> <li>2—I would have performed better than the automated vehicle in this situation</li> <li>3—In this situation, the automated vehicle performs well enough for me to engage in other activities (using my smartphone)</li> <li>4—The situation was risky</li> <li>5—The automated vehicle made an unsafe judgement in this situation</li> <li>6—The automated vehicle reacted appropriately to the environment</li> <li>7—In this situation, the behavior of the automated car surprised me</li> </ul>	Overall, how much do you trust the automated car?

Table 3. Trust scales used before, during, and after the experiment.

# 2.6. Data Collection and Analysis

Trust is a complex factor to assess: the tools that are used, their number (one or more), and the moment when trust is measured (before, during, or after interaction) varies across studies and may have an impact [38,39] on results. The variety of measurements in the scientific literature can be sub-divided into three categories: self-reported scales, behavioral measures, and physiological measures [11,12].

In this study, trust was measured using the model proposed in [40]. Based on our literature review, this model best matched the measurement of trust in the context of interaction with an automated vehicle at the time of our experimentation. Indeed, this model proposes various factors that would influence trust before and after interaction with the automated system.

Three different scales were selected to measure trust before, during, and after the experiment (Table 3). Different selection criteria were used to best assess the variability of trust dimensions among the three stages of interaction described in Hoff and Bashir's model [3]. However, in our results (Section 3) we distinguish global trust and multidimensional trust. On one side, global trust is measured using the same question across the three scales. This enables the evolution of trust to be measured from the beginning to the end of

the experiment. On the other side, multidimensional trust represents the dimensions of trust that differ before and during interaction with the system.

Initial and post-training trust was measured using the trust scale from [10]. As suggested by the model (Figures 6 and 7), this scale is made up of 24 questions measuring personality traits (questions 6 to 13) as well as pre-existing knowledge (questions 1 to 5 and 14 to 24). In the results section (Section 3), initial and post-training global trust refers to the answer to question 2 of this scale (*How much trust would you have in the automated car?*). Initial and post-training multidimensional trust scores are obtained by averaging the scores of all 24 questions (questions 5, 9, 13, 14, and 19 being negative; their scores were inverted before calculating the average).

During the experiment, trust was measured after each driving scenario using the scale taken from [36]. This scale was specifically designed as a first approach to empirically validate Hoff and Bashir's model [40]. Global trust is the answer to question 1 of this scale (*I trust the automated car in this situation*). The multidimensional trust score was obtained by averaging the scores of all 7 questions (questions 2, 4, 5, and 7 being negative; their scores were inverted before averaging—see Table 3).

Finally, trust was measured at the end of the experiment using the 1-item questionnaire from [37] (*Overall, how much trust do you have in the automated car?* See Table 3).

An initial and post-learning global trust score was computed based on item 2 (*How much trust would you have in the automated car?*) of the initial trust scale (adapted from [10]; see Appendix A). A global trust score was calculated for each driving scenario based on item 1 (*How well do you trust the automation in this situation?*) of the situational trust scale [36]. Finally, the post-experimental global trust was measured using a one-item survey [37] at the end of the experiment.



**Figure 6.** Factors that influence trust prior to interaction (initial trust in this paper) from Hoff and Bashir's model [40]. According to this model, prior to interaction with a system, user trust is influenced by his personality traits, age, gender, culture (dispositional trust), pre-existing knowledge of the technology (initial learned trust) which can be acquire through advertisement, peer recommendation or experience with similar system, and the situation in which the interaction is taking place (situational trust).



**Figure 7.** Factors that influence trust during interaction (measured after each scenario in this paper) from Hoff and Bashir's model [40]. During interaction, user trust is influenced by system design features and performance. The system performance is highly influenced by user reliance which is dynamically formed and adjusted during the interaction.

Perceived anthropomorphism is the degree to which users perceive humanness in a machine. In some studies, it has been measured through features that are uniquely human (agreeableness, openness, and civility) or typically human (extraversion, emotionality, warmth, and openness) [41,42]. Some studies explored sociability, human-likeness, and machine-likeness of robots [43,44]. Others explored the robot's consciousness, fakeness, and movement artificialness [45].

In this study, a self-reported scale from [14] was used to measure perceived anthropomorphism. It includes four questions which focused on the system's "mental" capacities. These questions asked participants how smart they thought the car was, how well the car could feel what was happening around it, how well it could anticipate what was about to happen, and how well it could plan a route. We chose this scale because we adopt the author's perspective on perceived anthropomorphism and because of the similarities between [14] and our experiment (an automated driving car with anthropomorphic features).

A perceived anthropomorphism score was calculated for each participant by averaging the answers to each of these four questions. The closer the score was to 10, the more the interface was perceived as anthropomorphic.

To test our hypotheses, we performed inferential tests. Quantitative data were analyzed using statistical tests adapted to the type of data. The statistical significance threshold has been set to the most common value in the field (*p*-value lesser than 0.05). These tests were combined with post hoc comparisons when further comparisons were needed (with *p*-value correction adjustments). Driving simulator data were quantitative (reaction time) while trust scale ratings (Likert scales) were combined to calculate a trust score. Statistical analyses have been performed on these scores.

# 3. Results

# 3.1. Perciption of Anthropomorphism

The hypothesis relating to Q1 was that perceived anthropomorphism will be higher when the interface integrates the visual assistant (holographic representation) compared to the interface integrating the voice assistant or the baseline interface.

Analysis of perceived anthropomorphism scores (Figure 8) did not support the hypothesis. Each interface received a relatively high perceived anthropomorphism score, between 7.5 and 8.5 out of 10. The vocal assistant interface had the highest score (8.5) compared to the visual assistant and baseline conditions (7.5). These differences were not statistically significant as shown by the Kruskal–Wallis test (chi squared = 0.99, df = 2, *p*-value = 0.61)



# Perceived\_Anthropomorphism

Interfaces

Figure 8. Perceived anthropomorphism score for each interface.

3.2. Trust

These analyses answer research question Q2. Our hypothesis was that trust levels are higher for participants experiencing the visual assistant than for those who have the voice assistant or the baseline. Table 4 presents the global trust scores obtained at each stage of the experiment.

**Table 4.** Average global trust score (interquartile ranges in parentheses), the \* sign indicates when the statistical test results are significant.

	Initial	Post-Training	Activation	<b>TOR 60</b>	TOR 10	Post-Experiment
Baseline	7 (1)	8 (0)	10 (1.25)	9.5 (1.25)	7 (3.5)	8 (2)
Vocal Assistant	6 (1)	7 (1.25)	8 (2.25)	9 (1.25)	8 (4.25)	7.5 (1.25)
Visual Assistant	6 (1)	7 (2)	8.5 (2)	9 (1.5)	7 (1.5)	7.5 (2)
Kruskal–Wallis test	0.0002 *	0.0284 *	0.0603	0.415	0.7895	0.6246

Results for the three driving scenarios (Activation, TOR 60, and TOR 10) show no significant differences among the three interfaces. However, initial and post-training global trust (Tables 5 and 6) are significantly different between the baseline and vocal assistant and between the baseline and visual assistant interfaces.

Table 5. Initial global trust.

	Visual Assistant	Vocal Assistant
Vocal Assistant	1.00	
Baseline	0.0005 *	0.0012 *

(*p*-values from the Mann–Whitney pair test using a Bonferroni correction; alt. hypothesis = Baseline > Assistant Vocal and Visual), the \* sign indicates when the statistical test results are significant.

Table 6. Post-training global trust.

	Visual Assistant	Vocal Assistant
Vocal Assistant	1.00	
Baseline	0.029 *	0.031 *

(*p*-values from the Mann–Whitney pair test using a Bonferroni correction; alt. hypothesis = Baseline > Assistant Vocal and Visual), the \* sign indicates when the statistical test results are significant.

Multidimensional trust scores showed no significant difference between the three interfaces except for the post-training score where participants had more trust with the baseline interface than with the two anthropomorphic interfaces (vocal and visual; see Tables 4 and 7). A two-by-two comparison of each interface confirmed this result (Table 8) for both global (Table 4) and multidimensional trust (Table 8).

**Table 7.** Multidimensional trust (median scores and interquartile range), the \* sign indicates when the statistical test results are significant.

	Initial	Post-Training	Activation	TOR60	TOR10
Baseline	6.7 (1.67)	8.03 (0.98)	8.43 (1.14)	9.07 (1.00)	5.71 (3.64)
Vocal Assistant	6.07 (1.68)	6.78 (1.70)	8.79 (1.64)	9.29 (1.39)	6.71 (2.96)
Visual Assistant	5.74 (1.05)	6.55 (1.54)	8 (1.25)	8.93 (1.11)	5.29 (2.61)
Kruskal–Wallis test	0.1271	0.0122 *	0.3823	0.5545	0.7003

Table 8. Multidimensional post-training trust.

	Visual Assistant	Vocal Assistant
Vocal Assistant Baseline	1.00 0.0068 *	0.0307 *
<i>n</i> -values from the Mann–Whitney pa	ir test using a Bonferroni correction;	alt. hypothesis = Baseline > Assistant

Vocal and Visual), the \* sign indicates when the statistical test results are significant.

Global and multidimensional trust scores were mostly equivalent, except for the initial trust measurement: multidimensional trust scores were not different across the three interface conditions (Table 5) as opposed to global trust score (Table 4). In summary, these results confirm our hypothesis only for the global trust score in the initial trust measurement.

# 3.3. Correlation between Anthropomorphism and Trust

The third research question (Q3; see Section 1) examined the relationship between perceived anthropomorphism and trust and hypothesized a correlation between them. A medium but significant correlation between perceived anthropomorphism and post-experimental global trust ( $\rho = 0.45$ , *p*-value < 0.05) supports this hypothesis (Figure 9).



**Figure 9.** Correlation ( $\rho = 0.45$ , *p*-value < 0.05) between perceived anthropomorphism (0: Not at all anthropomorphic—10 Very anthropomorphic, according to [14]) and post-experimental trust (according to the scale adapted from [37] (0: Not trustful—10 Very trustful).

## 3.4. Performance

The fourth question (Q4; see Section 1) was addressed by the hypothesis that, during TOR 60 and TOR 10 scenarios, driving performance (reaction time) is better when the takeover request is emitted by the visual assistant than by the voice assistant or the baseline interface.

Participant reactions were measured using SCANeR Studio data (reaction time, deviation from the axis, and time to collision) aggregated as a performance indicator. Analyses validated this hypothesis for the TOR 60 scenario.

# 3.4.1. TOR 60 Scenario

Here, participants received the takeover request alert 60 s before the end of the automated driving mode. In the voice and visual assistant conditions, the vocal message explained the cause of the takeover request (see Appendix A). This explanation was not available in the Baseline condition. Median participant reaction time (11.3 s) in the visual assistant condition was significantly shorter than in the vocal assistant or baseline conditions (Figure 10) as shown by the Kruskal–Wallis test (Chi-squared = 7.9595, df = 2, p-value = 0.0186).



#### Reaction time - TOR 60

Figure 10. Reaction time in TOR 60.

Thus, participants respond more quickly to the takeover request when it was emitted by the visual assistant compared to the baseline condition. These results were confirmed using a pairwise Mann–Whitney test with a Bonferroni correction (Table 9).

Table 9. Performance TOR 60, the \* sign indicates when the statistical test results are significant.

	Visual Assistant	Vocal Assistant
Vocal Assistant	0.7706	
Baseline	0.0068 *	0.0895

# 3.4.2. TOR 10 Scenario

In this scenario (see Figure 11), participants received a takeover request alert 10 s before the end of the automated driving mode. In the vocal assistant and visual assistant conditions, the assistant's voice message explains the cause of the takeover request (see Appendix A). This explanation is not available in the baseline condition. The hypothesis was that the assistant's explanation helped the participant better understand the situation and take over control more quickly in the voice and visual assistant conditions. Data did not support this hypothesis. Instead, reaction times were long (+7 s) in all groups and with no significant differences as shown by the Kruskal–Wallis test (chi-squared = 0.8631, df = 2, *p*-value = 0.6495).



Reaction time - TOR 10

Figure 11. Reaction time TOR 10.

### 4. Conclusions

## 4.1. Anthropomorphism

Anthropomorphic attributes such as voice or visual embodiment (voice and visual assistant conditions) added to a non-anthropomorphic interface (baseline condition) did not increase perceived anthropomorphism of the system, as was first hypothesized. Surprisingly, the three interfaces all received high anthropomorphism scores (around 7 on a scale of 10). Several explanations can be emphasized: the anthropomorphic attributes may not have reached the required quality to affect participants' perception. Indeed, improvements could be made to both the voice (tone variations) and the visual (animation) of the assistant. Alternatively, perhaps measured anthropomorphism is not influenced by the three interface types. The literature identifies two types of anthropomorphism [46,47]: conscious (mindful) and unconscious (mindless). The former is the ability to perceive human qualities in a machine, including its "intelligence" and the latter an individual's ability to view a machine as a "living entity". Our measurement did not distinguish these two types. More research could show whether anthropomorphic attributes might impact only one of these.

# 4.2. Trust

The hypothesis that higher perceived anthropomorphism led to higher trust was partially supported. Two scales measured trust (global and multidimensional) at the beginning of the experiment and at the end of each driving session. There was a moderate yet significant correlation between global trust and perceived anthropomorphism across participants regardless of the interface. This supports our hypothesis that perceived anthropomorphism and trust are positively correlated.

However, when comparing the three interface conditions across the experimental sessions, the pattern became more complex. Measurements revealed an initial difference in trust between the baseline and the two assistant conditions: the initial global trust of participants in the baseline group was significantly higher than that of the voice or visual assistant groups. There was no initial difference in the multidimensional trust score. At the end of the training session, both global and multidimensional trust were significantly higher in the Baseline group. All the metrics converged during the experimental trials when participants tested the interfaces in realistic conditions. At this stage, there were no significant differences among the three conditions regardless of the questionnaire used to assess trust.

We observed that the global trust level is the same for our three interfaces, baseline, voice, and visual assistant. According to our results, anthropomorphic interfaces cannot be recommended to car makers as a lever for increasing users' trust in driving automation. However, further research is needed to fully address this statement. Future research on this topic, may include an improved version of the virtual assistant, a larger sample group, and other factors contributing to users' trust such as acquired experience (driving scenarios participants are exposed to), and the system's level of transparency. Another focus of research could be how to best calibrate trust rather than simply increasing its level.

Another topic of interest is the discrepancy between the two trust assessment methods in which global trust scores seem more variable than multidimensional ones, especially for initial trust. Our results suggest that measuring trust with only one question could be less reliable than using a combination of questions and could lead to erroneous conclusions. Furthermore, since results obtained using the two methods seem to converge after the training session, it seems that participant's adaptation time and adjustment of trust levels should be considered when preparing the experimental protocol. Understanding this adaptation phenomenon could lead to further research in this work.

The hypothesis that higher anthropomorphism would positively impact driving performance was partially confirmed by the study of reaction time after a handover event in two scenarios (TOR 60 and TOR 10). The reaction time was significantly shorter in the visual assistant condition for a takeover request in 60 s. (TOR 60). For a 10 s takeover request (TOR 10), there was no significant difference in driving performance among the different interfaces. In that scenario, the addition of longer voice messages (vocal assistant) or of a hologram (visual assistant) did not affect reaction time.

This research partially supported our four hypotheses. There was a correlation between perceived anthropomorphism and trust. The anthropomorphic assistant, however, cannot be considered a better choice for improving these measures. All three interfaces, including the Baseline, received a high anthropomorphism score and, although initial trust was slightly higher in the baseline group, all groups reached the same level of trust after the training phase. As predicted, the visual and vocal assistant interfaces resulted in a faster takeover response but only in the 60 s takeover scenario and not in the more urgent 10 s one. These results are encouraging and raise exciting questions that require future research.

Author Contributions: All authors participated in the conceptualization, principally C.L.-G.; methodology: C.L.-G. and V.R. supervised by N.L., J.-M.A., B.L. and K.A.-F.; investigation and validation were performed by C.L.-G. supervised by N.L., J.-M.A. and K.A.-F.; formal analysis was performed by C.L.-G.; writing—original draft preparation was performed by K.A.-F. writing—review and editing, K.A.-F., N.L., C.L.-G., V.R. and J.-M.A., supervision was performed by K.A.-F., N.L. and J.-M.A.; funding acquisition was performed by IRT SystemX. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has been supported by the French government under the "France 2030" program, as part of the SystemX Technological Research Institute within the CMI project.

**Data Availability Statement:** Due to local privacy and ethical policy implementation, the results of this study can not be shared with the public.

Acknowledgments: This project is a collaborative project between manufacturers (Renault, Valeo, Saint-Gobain and Arkamys) and academic laboratories IMS-ENSC (Bordeaux) and CeRCA (Poitiers). Its objective is to explore human factors in the context of driving automation. It aims to define a multisensory interface (visual, auditory, and haptic) to accompany drivers in the specific use cases of transfer of control (gradient transition) and reassurance. We would like to thank Sabine LANGLOIS, Jean-Marc TISSOT, and Noé MONSAINGEON for their collaboration in this project.

**Conflicts of Interest:** The authors declare no conflict of interest. The sponsors had no role in the design, execution, interpretation, or writing of the study.

### Appendix A

Questionnaires used in the study.

In this experiment, participants took questionnaires in French. "Initial Trust" was kept in its original form in French and "situational trust", "perceived anthropomorphism" and "global trust" questionnaires were translated from their original form in English to French. Here, we present every questionnaire in English.

I—Initial Trust (adapted from [4]; English version)

Items 6 to 24 from [4]

Scale: Not at all (0) ----- Completely (10)

1—Do you think you know the automated car?

2—How much trust would you have in the automated car?

3—Do you think the automated car is useful?

4—Do you think the automated car is necessary?

5—Do you think that the automated car would interfere with your usual driving?

6—In everyday life, you tend to take risks

7—You tend to trust people

8—You believe that trusting someone means being able to trust them with something to do

9—You believe that it is necessary to know a person well to trust him/her

10—You think that the decision to trust someone depends on how you interact with them

11—You are generally suspicious of new technologies (cell phones, computers, internet, microwave ovens, etc.)

12—You think that new technologies are interesting

13—You think that new technologies are dangerous

What risks would you associate with using an automated car?

14—The risk of driving more dangerously

15—The risk of not knowing how to use the system

16—The risk of an accident with the car ahead

17-The risk of an accident with the car behind

18—The risk of being dependent on the system

19—The risk of losing the pleasure of driving

What benefits would you get from using an automated car?

20—Less stressful driving

21—Lightening the driving task (the automated car would allow me to perform a secondary activity such as reading or using my phone)

22—Easier driving task

23—Improved driving comfort

24—Safer driving

II—Situational Trust (adapted from [19])

Items 1 to 6 from [19]

Scale: Not at all (0) ----- Completely (10)

1—I trust the automation in this situation

2-I would have performed better than the automated vehicle in this situation

3—In this situation, the automated vehicle performs well enough for me to engage in other activities (such as using my smartphone) \*

4—The situation was risky

5—The automated vehicle made an unsafe judgement in this situation

6—The automated vehicle reacted appropriately to the environment

7—In this situation, the behavior of the automated vehicle (actions and decisions) surprised me \*\*.

\* The NDRT was updated to match the game used in our experiment. In the original questionnaire the task is "reading".

\*\* Item 7 was added by experimenters to include the dimension of surprise in trust evaluation. Here, we present the translated version from French to English.

III—Perceived Anthropomorphism—([7])

Scale: Not at all (0) ----- Completely (10)

1—How smart does this car seem?

2—How well do you think this car could perceive what is happening around it?

3—How well do you think this car could anticipate what is about to happen, before it actually happens?

4—How well do you think this car could plan the best route?

IV—Global Trust (adapted from [18])

Scale: Not at all (0) ----- Completely (10)

In general, do you trust the automated car? \*

\* The original version of this question mentioned adaptive cruise control. It was replaced by automated car in this experiment.

## Appendix B



Figure A1. Description of different levels of car automation by SAE International.

# References

- 1. Litman, T. Autonomous Vehicle Implementation Predictions: Implications for Transport Planning; Victoria Transport Policy Institute: Victoria, BC, Canada, 2018; Volume 28.
- 2. Biondi, F.; Alvarez, I.; Jeong, K.-A. Human–Vehicle Cooperation in Automated Driving: A Multidisciplinary Review and Appraisal. *Int. J. Hum.–Comput. Interact.* **2019**, *35*, 932–946. [CrossRef]
- Biondi, F.; Lohani, M.; Hopman, R.; Mills, S.; Cooper, J.; Strayer, D. 80 MPH and out-of-the-loop: Effects of real-world semiautomated driving on driver workload and arousal. In Proceedings of the Conference of the Human Factors and Ergonomics Society, Los Angeles, CA, USA, 1–5 October 2018.
- 4. Stanton, N.A.; Young, M.S. Driver behaviour with adaptive cruise control. Ergonomics 2005, 48, 1294–1313. [CrossRef] [PubMed]
- 5. Vollrath, M.; Schleicher, S.; Gelau, C. The influence of cruise control and adaptive cruise control on driving behaviour—A driving simulator study. *Accid. Anal. Prev.* 2011, 43, 1134–1139. [CrossRef]
- Liu, P.; Yang, R.; Xu, Z. Public Acceptance of Fully Automated Driving: Effects of Social Trust and Risk/Benefit Perceptions. *Risk* Anal. 2019, 39, 326–341. [CrossRef] [PubMed]
- 7. Knack, S. Trust, Associational Life and Economic Performance. In *The Contribution of Human and Social Capital to Sustained Economic Growth and Well-Being*; Helliwell, J.F., Ed.; Human Resources Development: Quebec, QC, Canada, 2001.
- Mayer, R.C.; Davis, J.H.; Schoorman, F.D. An Integrative Model of Organizational Trust. Acad. Manag. Rev. 1995, 20, 709–734. [CrossRef]
- 9. Muir, B.M. Trust in automation: I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics* **1994**, *37*, 1905–1922. [CrossRef]
- 10. Rajaonah, B. Rôle de la Confiance de L'opérateur dans son Interaction avec une Machine Autonome sur la Coopération Humain-Machine. Doctor's Thesis, Paris 8, Paris, French, 2006.
- 11. French, B.; Duenser, A.; Heathcote, A. *Trust in Automation A Literature Review*; Commonwealth Scientific and Industrial Research Organisation (CSIRO): Canberra, Australia, 2018.
- 12. Kohn, S.C.; de Visser, E.J.; Wiese, E.; Lee, Y.-C.; Shaw, T.H. Measurement of Trust in Automation: A Narrative Review and Reference Guide. *Front. Psychol.* **2021**, *12*, 604977. [CrossRef]
- 13. Jian, J.-Y.; Bisantz, A.M.; Drury, C.G. Foundations for an Empirically Determined Scale of Trust in Automated Systems. *Int. J. Cogn. Ergon.* 2000, *4*, 53–71. [CrossRef]
- 14. Waytz, A.; Heafner, J.; Epley, N. The Mind in the Machine: Anthropomorphism Increases Trust in an Autonomous Vehicle. *J. Exp. Soc. Psychol.* **2014**, *52*, 113–117. [CrossRef]
- 15. Garcia, D.; Kreutzer, C.; Badillo-Urquiola, K.; Mouloua, M. Measuring Trust of Autonomous Vehicles: A Development and Validation Study. In *HCI International 2015—Posters' Extended Abstracts*; Stephanidis, C., Ed.; Springer International Publishing: Cham, Switzerland, 2015; Volume 529, pp. 610–615. [CrossRef]
- 16. Yagoda, R.E.; Gillan, D.J. You want me to trust a ROBOT? The development of a human–robot interaction trust scale. *Int. J. Soc. Robot.* **2012**, *4*, 235–248. [CrossRef]
- 17. Hester, M.; Lee, K.; Dyre, B.P. "Driver Take Over": A preliminary exploration of driver trust and performance in autonomous vehicles. In *Human Factors and Ergonomics Society Annual Meeting*; Sage: Los Angeles, CA, USA, 2017; Volume 61, pp. 1969–1973.
- 18. Gaudiello, I.; Zibetti, E.; Lefort, S.; Chetouani, M.; Ivaldi, S. Trust as indicator of robot functional and social acceptance. An experimental study on user conformation to iCub answers. *Comput. Hum. Behav.* **2016**, *61*, 633–655. [CrossRef]
- 19. Hergeth, S.; Lorenz, L.; Vilimek, R.; Krems, J.F. Keep your scanners peeled: Gaze behavior as a measure of automation trust during highly automated driving. *Hum. Factors J. Hum. Factors Ergon. Soc.* **2016**, *58*, 509–519. [CrossRef] [PubMed]
- 20. Morris, D.M.; Erno, J.M.; Pilcher, J.J. Electrodermal response and automation trust during simulated self-driving car use. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 2017, *61*, 1759–1762. [CrossRef]
- Miller, D.; Johns, M.; Mok, B.; Gowda, N.; Sirkin, D.; Lee, K.; Ju, W. Behavioral measurement of trust in automation. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* 2016, 60, 1849–1853. [CrossRef]
- 22. Lee, J.D.; See, K.A. Trust in Automation: Designing for Appropriate Reliance. Hum. Factors 2004, 46, 50–80. [CrossRef] [PubMed]
- 23. Parasuraman, R.; Riley, V. Humans and Automation: Use, Misuse, Disuse, Abuse. Hum. Factors 1997, 39, 230-253. [CrossRef]
- 24. Niu, D.; Terken, J.; Eggen, B. Anthropomorphizing information to enhance trust in autonomous vehicles. *Hum. Factors Ergon. Manuf. Serv. Ind.* 2017, *28*, 352–359. [CrossRef]
- 25. Large, D.R.; Harrington, K.; Burnett, G.; Luton, J.; Thomas, P.; Bennett, P. To Please in a Pod: Employing an Anthropomorphic Agent-Interlocutor to Enhance Trust and User Experience in an Autonomous, Self-Driving Vehicle. In Proceedings of the Automotive UI '19: 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Utrecht, The Netherlands, 21–25 September 2019; pp. 49–59.
- 26. Reeves, B.; Nass, C. The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Pla. In *Bibliovault OAI Repository*; the University of Chicago Press: Chicago, IL, USA, 1996.
- 27. Valette-Florence, R.; de Barnier, V. Les lecteurs sont-ils capables d'anthropomorphiser leur magazine? Une réponse par la méthode de triangulation. *Manag. Avenir* **2009**, 27, 54–72. [CrossRef]
- Admoni, H.; Scassellati, B. A Multi-Category Theory of Intention. CogSci, 2012. Available online: https://scazlab.yale.edu/sites/ default/files/files/Admoni-Cogsci-12.pdf (accessed on 19 October 2023).

- 29. Epley, N.; Waytz, A.; Cacioppo, J.T. On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychol. Rev.* 2007, 114, 864–886. [CrossRef]
- Złotowski, J.; Strasser, E.; Bartneck, C. Dimensions of anthropomorphism: From humanness to human likeness. In Proceedings of the HRI'14: ACM/IEEE International Conference on Human-Robot Interaction, Bielefeld, Germany, 3–6 March 2014; pp. 66–73.
- DiSalvo, C.F.; Gemperle, F.; Forlizzi, J.; Kiesler, S. All robots are not created equal: The design and perception of humanoid robot heads. In Proceedings of the DIS02: Designing Interactive Systems 2002, London, UK, 25–28 June 2002; pp. 321–326.
- 32. Fink, J. Anthropomorphism and Human Likeness in the Design of Robots and Human-Robot Interaction. In *Social Robotics*; Ge, S.S., Khatib, O., Cabibihan, J.-J., Simmons, R., Williams, M.-A., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 199–208. [CrossRef]
- Häuslschmid, R.; von Bülow, M.; Pfleging, B.; Butz, A. SupportingTrust in Autonomous Driving. In Proceedings of the IUI'17: 22nd International Conference on Intelligent User Interfaces, Limassol, Cyprus, 13–16 March 2017; pp. 319–329.
- Zihsler, J.; Hock, P.; Walch, M.; Dzuba, K.; Schwager, D.; Szauer, P.; Rukzio, E. Carvatar: Increasing Trust in Highly-Automated Driving Through Social Cues. In Proceedings of the AutomotiveUI'16: 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Ann Arbor, MI, USA, 24–26 October 2016; pp. 9–14.
- 35. Vermersch, P. L'entretien d'explicitation en Formation Initiale et Continue; ESF: Paris, France, 1994.
- Holthausen, B.E.; Wintersberger, P.; Walker, B.N.; Riener, A. Situational Trust Scale for Automated Driving (STS-AD): Development and Initial Validation. In Proceedings of the AutomotiveUI '20: 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, Virtual Event, 21–22 September 2020.
- Cahour, B.; Forzy, J.-F. Does projection into use improve trust and exploration? An example with a cruise control system. *Saf. Sci.* 2009, 47, 1260–1270. [CrossRef]
- 38. Hancock, P.A.; Billings, D.R.; Schaefer, K.E.; Chen, J.Y.C.; de Visser, E.J.; Parasuraman, R. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. *Hum. Factors* 2011, *53*, 517–527. [CrossRef]
- Schaefer, K.E.; Billings, D.R.; Szalma, J.L.; Adams, J.K.; Sanders, T.L.; Chen, J.Y.; Hancock, P.A. A Meta-Analysis of Factors Influencing the Development of Trust in Automation: Implications for Human-Robot Interaction (ARL-TR-6984). Army Research Lab Aberdeen Proving Ground MD Human Research and Engineering Directorate. Available online: https://apps.dtic.mil/docs/ citations/ADA607926 (accessed on 1 December 2020).
- 40. Hoff, K.A.; Bashir, M. Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust. *Hum. Factors J. Hum. Factors Ergon. Soc.* 2015, *57*, 407–434. [CrossRef]
- 41. Haslam, N.; Loughnan, S.; Kashima, Y.; Bain, P. Attributing and denying humanness to others. *Eur. Rev. Soc. Psychol.* 2008, 19, 55–85. [CrossRef]
- Moussawi, S.; Koufaris, M. Perceived Intelligence and Perceived Anthropomorphism of Personal Intelligent Agents: Scale Development and Validation. In Proceedings of the Hawaii International Conference on System Sciences, Maui, HI, USA, 8–11 January 2019.
- 43. Kiesler, S.; Powers, A.; Fussell, S.R.; Torrey, C. Anthropomorphic interactions with a robot and robot-like agent. *Soc. Cogn.* **2008**, 26, 169–181. [CrossRef]
- 44. Powers, A.; Kiesler, S. The advisor robot: Tracing people's mental model from a robot's physical attributes. In Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction, Salt Lake City, UT, USA, 2–3 March 2006.
- 45. Bartneck, C.; Kulić, D.; Croft, E.; Zoghbi, S. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *Int. J. Soc. Robot.* **2009**, *1*, 71–81. [CrossRef]
- 46. Kim, Y.; Sundar, S.S. Anthropomorphism of computers: Is it mindful or mindless? *Comput. Hum. Behav.* 2012, 28, 241–250. [CrossRef]
- 47. Esfahani, M.S.; Reynolds, N.; Ashleigh, M. Mindful and mindless anthropomorphism: How to facilitate consumer comprehension towards new products. *Int. J. Innov. Technol. Manag.* **2020**, *17*, 2050016. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.