“I will raise my hand and say ‘I over-trust Autopilot’. I use it too liberally” Drivers’ reflections on their use of partial driving automation, trust, and perceived safety

Nordhoff, Sina; Hagenzieker, Marjan

DOI
10.1016/j.trf.2024.09.021

Publication date
2024

Document Version
Final published version

Published in
Transportation Research Part F: Traffic Psychology and Behaviour

Transportation Research Part F: Psychology and Behaviour 107 (2024) 1105–1124

Contents lists available at ScienceDirect

Transportation Research Part F: Psychology and Behaviour

journal homepage: www.elsevier.com/locate/trf

"I will raise my hand and say 'I over-trust Autopilot'. I use it too liberally" – Drivers' reflections on their use of partial driving automation, trust, and perceived safety

Sina Nordhoff[a,b,*], Marjan Hagenzieker[a]

[a] Department Transport & Planning, Delft University of Technology, the Netherlands
[b] Electric Vehicle Research Center, Institute of Transportation Studies, University of California, Davis, United States

A R T I C L E I N F O

Keywords:
Partial driving automation
Trust
Perceived safety
Natural Language Processing (NLP)
Automation use

A B S T R A C T

Introduction: Partially automated cars are on the road. Trust in automation and perceived safety are critical factors determining use of automation. **Background:** Drivers misuse partially automated driving systems. Misuse is associated with mis-calibrated trust in the automation. **Research gap:** Little is known about the factors impacting the perceived safety when using partial driving automation. **Research objective:** The main objective of the present study is to provide a comprehensive driver perspective on the psychological aspects of automation use pertaining to trust in automation, perceived safety, and its relationship with use of automation. **Method:** Semi-structured interviews (n = 103) were conducted with users of partially automated driving systems. Supplemented with content analysis, natural language processing (NLP) techniques were applied to perform automatic text processing. Guided seed-term analysis was conducted to identify the number of occurrences of the subcategories in the dataset. **Main results:** We identified human operator-related, automation-related, and environmental factors of trust and perceived safety. The identified factors were more strongly associated with perceived safety than with trust. Participants with physical and visual impairments reported to feel safer using the automation compared to driving manually. Neurotic behavior during manual driving contributed to lower trust and perceived safety using the automation. A correct mental model of the capabilities and limitations of the automation did not guarantee proper automation use. A novel conceptual, process-oriented model, titled PTS-a (predicting trust in and perceived safety of automation use), synthesizes the results of the data analysis. Informed by the cognition-leads-to-emotions approach, the model posits that trust as cognition precedes perceived safety as affective construct. Trust and perceived safety determine how human operators (mis-, dis-)use the automation. **Future research:** We recommend future research to perform experimental studies to identify cognitive-related thoughts and beliefs pertaining to trust in automation and perceived safety to contribute to the operationalization of these constructs, and unravel the nature of their relationship.

* Corresponding author.
E-mail address: snordhoff@ucdavis.edu (S. Nordhoff).

https://doi.org/10.1016/j.trf.2024.09.021
Received 13 February 2024; Received in revised form 19 May 2024; Accepted 25 September 2024
Available online 11 November 2024

1. Introduction

Partially automated vehicles are on the road. Trust and perceived safety have been identified as key factors predicting acceptance and use of automation (Ljubi & Groznik, 2023). There is a long research tradition investigating trust in automation (Hancock et al., 2011; Lee & See, 2004; Schaefer et al., 2014). Hancock et al. (2011) and Schaefer, Chen, Szalma, & Hancock (2016) developed a framework for the conceptual organization of trust, which includes three main elements, including trust calibration, trust outcomes, and trust development. Trust calibration concerns the process when users have inappropriate levels of trust in the automation that is not calibrated to the capabilities of the automation, which can have severe negative consequences. In cases of overtrust, overreliance or misuse of the automation may follow (Lee & See, 2004; Muir, 1994; Parasuraman & Riley, 1997), while disuse of the automation may follow in cases of too little trust in the automation (i.e., distrust) (Hancock et al., 2011; Lee & See, 2004). Overtrust is defined as "poor calibration in which trust exceeds system capabilities", while in cases of distrust "trust falls short of the automation's capabilities" (Lee & See, 2004) (p. 55). An appropriate level of trust is defined as "level of trust that is "reasonable" in relation to the actual performance of the AD system" (automated driving system) (Ekman, Johansson, & Sochor, 2017) (p. 95). Studies revealed that users of partially automated vehicles misuse the automation (Banks, Eriksson, O'Donoghue, & Stanton, 2018; Kim, Song, & Doerzaph, 2021; Mueller, Cicchino, & Calvanelli, 2023; Nordhoff et al., 2023).

Mueller et al. (2023) argue that if drivers trust the automation too much, their ability to intervene rapidly enough may be impaired, leading to uncomfortable or dangerous situations. Similarly, we posit that perceived safety can contribute to improper automation use if users feel too safe using the automation, and engage in potentially unsafe driving behaviors, limiting their capability to respond to objects and events in the driving environment as requested by the partially automated driving system. Conversely, low perceived safety may contribute to disusing or disengaging the automation, which may undermine the expected benefits of partial driving automation (Mueller, Reagan, & Cicchino, 2021). Perceived safety has been strongly related to trust in previous studies (He, Stapel, Wang, & Happee, 2022). Trust is defined as the human operator's attitude that the automation will help the human operator to achieve his / her goals in a situation that is characterized by uncertainty and vulnerability (Chiou & Lee, 2023; Lee & See, 2004). Perceived safety is related to a person's subjective feelings, such as relaxation, and perceived risk (Xu et al., 2018).

The trust development process is concerned with identifying the factors influencing trust. These can be grouped in factors related to the human, automation, environment, and training (Billings, Schaefer, Chen, & Hancock, 2012; Hancock et al., 2011; Schaefer et al., 2014; Schaefer et al., 2016). We will only report the factors that are relevant for the context of the present study.

1.1. *Human operator-related factors*

Human-related factors can be broadly categorized into human traits, human states, cognitive and emotive factors. Factors representing human traits encompass the age, gender, ethnicity, or personality of human operators (Schaefer et al., 2016). Trust propensity is defined as trait-based or dispositional trust that describes an individual's general inclination to be trusting of others (Hoff & Bashir, 2015; Schaefer, Perelman, Gremillion, Marathe, & Metcalfe, 2021). Driving-related personality is typically represented by the personality constructs locus of control, sensation seeking, and neuroticism. People with an internal locus of control (i.e., personal tendency to maximize the possibility of positive and minimize the possibility of negative events) have been shown to be less likely to trust the automation, while people with an external locus of control (i.e., lower tendency to not accept personal responsibility for the outcome of events) are more likely to (over-) trust the automation (Rudin-Brown & Noy, 2002). High-sensation seekers (i.e., individuals with a higher tendency to seek novel, varied, complex, and intense sensations and experiences, and the willingness to take risks to make these experiences) (Zuckerman, 1994) have been shown to be more likely to engage in risky driving (Jonah, 1997) and develop (over-)trust in the automation than low-sensation seekers (Rudin-Brown & Noy, 2002). A positive bivariate correlation between sensation seeking and trust in automation has been reported in Palmeiro, van der Kint, Hagenzieker, Van Schagen, and De Winter (2018) and Zhang et al. (2020). Neuroticism (i.e., low emotional stability and impulse control, associated with negative emotional states, such as anxiety, guilt, and fear, Mooradian, Renzl, and Matzler (2006)) has been associated with lower trust in the automation (Zhang et al., 2020). These studies have offered scientific evidence about how a person's driving-related personality influences trust in the automation. Little is known about how these personality traits affect the perceived safety of using partial driving automation. One of the few relevant studies on Human-Robot-Interaction (HRI) has revealed a negative effect of neuroticism on perceived safety (Akalin, Kristoffersson, & Loutfi, 2022).

Cognitive factors capture the human operator's understanding of how the automation works, ability to use the automation, and expectancy of the automation. The aspect 'understanding how the automation works' is captured by a human operator's experience with the automation (Schaefer et al., 2016). The notion of 'learned trust' is trust that is based on experience (Hoff & Bashir, 2015). Experience with the automation has contributed to an increase in trust in the automation, while the effect of experience on perceived safety was not examined by the authors in the same study (Xu et al., 2018). A related construct is subject matter expertise, which is defined as an understanding of the subject matter that develops from experience. Individuals with subject matter expertise were less likely to rely on the automation than novice users in domains other than automated driving (Hoff & Bashir, 2015). Experience is also referred to as familiarity (Wicki, Brückmann, Quoss, & Bernauer, 2023). Another related construct is knowledge. Knowledge refers to how much individuals know about and can memorize about a topic (Wicki et al., 2023). Knowledge from past interactions with the automation contributed to trust (Hoff & Bashir, 2013). Studies examining the acceptance of electric vehicles have shown that knowledge led to higher acceptance (e.g., Wicki et al., 2023).

1.2. Automation-related factors

Automation-related factors can be divided into factors representing the capability and features of the automation (Schaefer et al., 2016). Reliability and predictability are capability-based factors influencing trust in the automation (Hancock et al., 2011; Hoff & Bashir, 2015; Lee & Moray, 1992; Rudin-Brown & Noy, 2002). Reliability is concerned with the consistency of the automation (Hoff & Bashir, 2015), which develops from accumulated knowledge over time (Johnson-George & Swap, 1982). Predictability refers to the extent to with an operator can anticipate or foresee future behavior of the automation (Muir, 1994) based on the extent to which the automation performs in a way that is consistent with the human operator's expectations (Hoff & Bashir, 2015). Perceived predictability is determined by the consistency and desirability of the automation behavior, and can be assessed by the actual predictability and observability of the automation behavior (Muir, 1994; Rempel, Holmes, & Zanna, 1985).

Transparency and explainability are feature-based automation-related factors critical for appropriate levels of trust (Endsley, 2016). They address the understandability and predictability of the automation's intent, future plans, and process of reasoning (Chen et al., 2014; Endsley, 2016). Both can improve the observability of the automation behavior by making its behavior transparent and easy to understand (Endsley, Bolté, & Jones, 2003; Muir, 1994). Maarten Schraagen et al. (2021) have shown that providing participants with explanations about the behavior of a partially automated driving system resulted in the lowest levels of trust, whereas trust was higher when the design of the interface in the car offered transparent information. In a study investigating Human-Robot-Interaction (HRI), unpredictable robot behavior did not influence perceived safety. On the other hand, transparency about the reasons for the robot's behavior positively influenced perceived safety (Akalin et al., 2022).

1.3. Environmental factors

Environmental factors include factors in the physical environment and advertising, training, and education. The notion of situational trust reflects the context-specific nature of trust, with the external environment influencing trust (Hoff & Bashir, 2013). The stability of the external environment is important for the predictability of the automation (Muir, 1994; Rempel et al., 1985), and thus trust. Factors in the physical environment influencing trust include, but are not limited to, weather conditions and road infrastructure and design characteristics, as well as the behavior of the automation in the environment (Stocker, 2022; Walker et al., 2023). We previously found that weather conditions as well as road infrastructure and design characteristics contributed to disengaging the partial automated driving system (Nordhoff, 2024).

Media, advertising, and training represent other external factors in the environment that are considered in the present study. They shape driver's expectations about how these automated driving systems function. It is still unclear how drivers' expectations will influence the use and adoption of partial driving automation (Russell, Atwood, & McLaughlin, 2021). The capabilities of the partial driving automation have shown to be oversold by media and marketing, increasing the likelihood of overtrust and improper use of the automation (Dixon, 2020), and the development of unrealistic expectations about the actual capabilities of the automation (Russell et al., 2021). Training can improve driver's knowledge about the system functionality, helping drivers to familiarize themselves with the functionality, and correct behavior, contributing to the creation of appropriate trust in the automation (Ekman et al., 2017).

1.4. The present study

The present study posits that understanding the role of perceived safety for proper automation use requires the identification of the influencing factors of perceived safety in relation to trust. Most of the studies on trust and perceived safety are simulator or questionnaire studies involving participants without direct, continued experience with partially automated driving systems. This makes the identification of the factors influencing trust and perceived safety challenging (He et al., 2022). Only a few studies have examined the factors influencing trust in partial driving automation (Hardman et al., 2021), and even less is known about the factors influencing perceived safety (Cao et al., 2021). A comprehensive driver perspective on the psychological aspects of automation use pertaining to trust, perceived safety, and automation use and their relationships is missing. To address these gaps, semi-structured interviews were conducted with users of vehicles equipped with partial driving automation to address the following two research questions:

1) What are human-operator-related, automation-related-, and environmental factors influencing trust and perceived safety?
2) To what extent does trust in automation influence perceived safety, and use of partial driving automation?

2. Method

2.1. Recruitment

The study was approved by the Human Research Ethics Committee of Delft University of Technology (ID: 2316). We applied a two-stage recruiting. In the fist stage, we applied convenience sampling to recruit participants in social media forums via special interest groups that we identified before (e.g., Reddit, Facebook, Discord, Twitter). In the second sage, we used snowball sampling to contact additional people via email that participants recruited in the first stage referred to as potential participants.

Participants with access to FSD Beta received information by Tesla via email on how to use the system prior to first use. The email informed participants that Full Self-Driving is in a limited early access Beta phase, and must be used with additional caution because it may do the wrong thing at the worst time. Participants were instructed to always keep their hands on the wheel, pay extra attention to the road and not become complacent. Participants were informed that they are responsible for remaining alert and active and must be prepared to take action at any time.

2.2. Procedure

The interviews were conducted via Zoom. A link to an interview protocol created on Qualtrics (www.Qualtrics.com) was shared with participants at the beginning of the interview to reduce the subjectivity of interview research and minimize the intervention of the interviewer.

At the start of the interview, participants were asked to provide their informed consent to participate in the study. With the first few questions (Q8–Q11), participants were presented four general statements about the operation of Autopilot and FSD Beta pertaining to whether Autopilot makes driving autonomous (Q8); whether there were safety issues with Autopilot (Q9); whether Autopilot was a hands-free feature (Q10), and whether FSD Beta was safer than a human (Q11) to be answered on a scale from 1 (Yes), 2 (No), to 3 (I don't know). Second, participants were asked to indicate to what extent they felt safe when Autopilot and FSD Beta were active, and to motivate this decision (Q12). Third, they were asked what or how they felt when they felt safe (Q13), and what it was about Autopilot and FSD Beta that made them feel safe or unsafe (Q14). They were also asked to remember the situations in which they typically felt unsafe when Autopilot and FSD Beta were active (Q15), and to explain what Autopilot and FSD Beta can do to support their safety (Q16). Next, they indicated to what extent feeling safe or feeling unsafe impacted how they used Autopilot and FSD Beta on their next drives or in the future (Q17), and whether their perceived safety changed over time (Q18). We also asked participants to rate their trust in Autopilot and FSD Beta (Q19), to explore what Autopilot and FSD Beta could do to support their trust (Q20), whether their trust or distrust in Autopilot and FSD Beta impacted how they used Autopilot and FSD Beta on their next drives / in the future (Q21), and whether their trust changed over time, and if so how (Q22).

This first interview part consisted of several closed- and open-ended questions (Q1–Q35). In the second part, participants were presented with personal questions pertaining to their socio-demographic profile, travel behavior (e.g., age, gender, highest level of education, frequency of use of Autopilot and FSD Beta), and general attitudes towards traffic safety. The interviewer left the interview at the end of the first part to give participants the anonymity to complete the personal questions on their own.

Table A1 in the Appendix provides an overview of all closed-ended questions that were asked during the interview. We will only present the responses to the questions subjected to the analysis of the present study as the remaining questions were addressed in previous studies (Nordhoff, 2024; Nordhoff et al., 2023).

2.3. Data analysis

The data analysis was performed in six steps.

First, the interviews were transcribed verbatim.

Second, content analysis was conducted to identify human-operator-, automation-, and environment-related factors influencing trust and perceived safety. The factors were inductively developed from the data using common text analysis methods, such as writing notes, searching for keywords, and jumping between text passages. The development of the factors was derived from repetitions, similarities, and differences of key words and phrases. The subcategories were clearly defined to be operationalized in future studies, and understandable to experts and laymen. As the subcategories were inductively developed from the data, they were directly applicable to the data (Glaser, Strauss, & Strutzel, 1968). The development of the subcategories was iterative and emergent as the researcher revisited the transcripts again, comparing the subcategories with the literature and consequently refining them. This part of the analysis was conducted in Atlas.ti (Version 22.0.2). A subcategory had to be mentioned by at least five participants to qualify as subcategory.

Third, these subcategories were closer inspected using natural language processing techniques. This step involved data pre-processing conducted in several steps in a specific order in line with Maier et al. (2021). First, word tokenization is applied, which transforms sentences into token-sized pieces or individual words (Thushari et al., 2023). After tokenization, capital letters were converted to lowercase. Next, punctuation and special and spurious characters (e.g., periods, commas), words with 2 or less than 2, or more than 30 digits as well as numbers were removed. Next, stop-words and additional words not carrying substantial meaning or misspelled words were removed (e.g., 'think', 'guess', 'yes').

Fourth, to identify the occurrences of the subcategories in the data, seed terms were identified. The identification of the seed terms was based on theory and data, which represents a common research practice (Watanabe & Zhou, 2022). To improve the process of the identification of the seed terms, the global vectors model GloVe by Pennington, Socher, and Manning (2014) was applied to identify synonyms of the seed terms as in Ploessl, Just, and Wehrheim (2021). We used a large database of web text titled Common Crawl suitable for our purpose containing 300-dimensional word vectors for 600.000 words to identify 15 additional seed terms per subcategory (Pennington, Socher, &, Manning, n. d.). To further enrich our list of seed terms using laymen-friendly language, we scraped the social media platform Reddit, analyzing 998 posts in the subreddit 'SelfDrivingCars'. Finally, we inspected the word collocations of our subcategories, i.e., the words that commonly co-occurred together, to further enrich our list of seed terms. The list of seed terms generated in this process was discussed between the authors, and seeds that were not directly applicable were omitted.

Fifth, to count the occurrence of the subcategories, it was assumed that each subcategory is represented in a window size of 10, i.e., in the 10 words before and the 10 words after each seed term representing a subcategory. A subcategory was assigned a frequency of 1 if at least two seed terms representing a subcategory were mentioned once in this defined window size. The total number of mentions of a subcategory equaled the total number of occurrences of at least two seed terms capturing the subcategory in the defined window size across all interview transcripts. Steps 3–5 were conducted in Python.

Sixth, illustrative quotes from participants were selected to portray the meaning of the categories. Multiple mentions of a subcategory per respondent were not discarded but merged with the other mentions of the subcategory by the respondent. As a result, some of the quotes represent clusters of sentences mentioned by the same participant at different points in time during the interview.

Finally, the grounded theory approach was applied to develop a conceptual model (Glaser et al., 1968; Jabareen, 2009). This step included categorizing the categories, naming, and describing each category, its role in the model as well as presenting its references. Next, categories were integrated, which involved grouping similar categories together. In the final step, the categories were synthesized into a conceptual model, which is an iterative process, involving a repetitive synthesis and resynthesis until we recognized a conceptual model that made sense to us (Jabareen, 2009). As the model links the categories, it provides important impetus for future studies for the development and testing of hypotheses pertaining to the relationships between contextual factors and trust, perceived safety, and use of partial driving automation.

3. Results

3.1. Participants

In total, we performed 103 semi-structured interviews between Feburary and June 2022. One interview lasted on average 01:18:05 h. Participants were on average 43 years old (SD = 14 years). 91 % of our participants were Male, and 9 % were Female. 52 % had a Bachelor's or Master's degree, 27 % reported to have a college degree, 13 % a high school diploma, and 8 % had a PhD degree. They predominantly resided in California (20 %), Colorado (8 %), and Florida (7 %). Respondents reported to be engineers (30 %), managers (8 %), or they were retired (7 %). 82 % of participants used both Autopilot and FSD Beta, while 18 % of participants only had access to Autopilot. The average use of Autopilot and FSD Beta was 26.8 and 8.14 months, respectively. Participants made almost daily use of Autopilot and FSD Beta as indicated by the mean value of 4.11 for Autopilot and 4.50 for FSD Beta, measuring the reported frequency of use on a scale from 1 = less than monthly, 2 = less than weekly but more than once a month, 3 = 1–2 times a week, 4 = 3–4 times a week to 5 = at least five times a week.

3.2. Knowledge about operation of Autopilot and FSD Beta

We analysed participants' knowledge about the operation of Autopilot and FSD Beta by four main questions. The question 'The current Autopilot does make driving autonomous. Is that correct?' was answered with 'No' by 79 % of participants, indicating that participants were generally knowledgeable about Autopilot not making driving autonomous, while 17 % of respondents selected 'Yes', and 3 % were undecided (selected the response option 'I don't know').

„It is a Level 1, Level 2 system, so you are in control. You are responsible for what the car is doing. You are expected to take control." (R001)
„That's the key that it is a Level 2 autonomous system, which means the human always has to be aware what the car is doing." (R079)

75 % of participants believed that were no safety issues with Autopilot (selected the response option 'Yes' to the question 'There are no safety issues with Autopilot. Is that correct?'), while 14 % of participants said 'No' to the question and 11 % 'I don't know'.

„Safety issues. It's not perfect, and you do need to pay attention and watch it. You have to babysit the car while it's on Autopilot." (R016)
„Autopilot, there's no issues, no problems. It does exactly what it's supposed to do. There's no issues, no problems, very happy with it." (R069)

Autopilot was considered a hands-on feature by 90 % of respondents, while 8 % selected the response option 'No' to the question 'Autopilot is a hands-on feature. Is that correct?', and 2 % did not know, picking the response option 'I don't know'.

„Yes, Autopilot is a hands-on feature. It could probably be a hands-off feature on the Interstate. Autopilot's mistakes on the highway are slow enough so that you could get your hands up in time, and grab the steering wheel and prevent an accident. I only put my hands on the highway to appease the system." (R007)
„It's most definitely a hands-on feature mainly also because the car requires that you pay attention. It's an assistance feature. We are not anything anywhere near an autonomous driving system." (R016)

Finally, 63 % did not consider FSD Beta safer than a human, selecting the response option 'No' to the question 'Tesla FSD Beta is safer than a human. Is that correct?', while 22 % picked 'Yes', and 16 % chose the option 'I don't know'.

„It's not even safer than a human with a human with their hands on the steering wheel, ready to take over." (R007)
„Absolutely not safer than a human. It's more distracted to things it doesn't need to get distracted with." (R071)

3.3. Trust in Autopilot and FSD Beta

36 % of participants did not trust FSD Beta compared to 0 % of participants reporting to distrust Autopilot. 93 % of participants indicated to trust Autopilot, 55 % of which reported to trust Autopilot a lot, as demonstrated by the following participant:

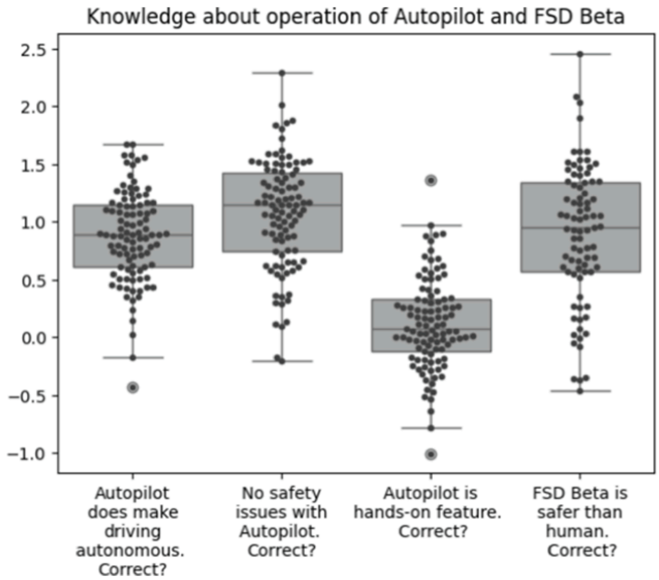

Fig. 1a. Box plots of responses for questions about knowledge about operation of Autopilot & FSD Beta measured on a scale from 1 = Yes, 2 = No, to 3 = I don't know.

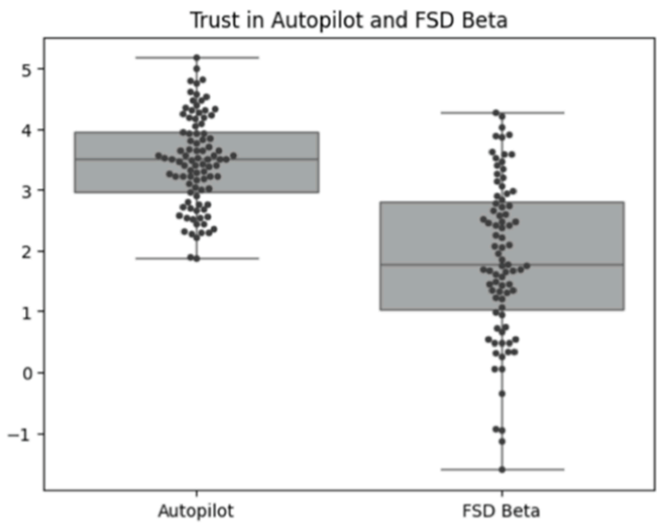

Fig. 1b. Box plots of responses for trust in Autopilot and FSD Beta measured on a scale from 1 = I don't trust it at all, 2 = I don't trust it, 3 = I neither don't trust it at all nor trust it a lot, 4 = I trust it at all, to 5 = I trust it a lot.

> „*Autopilot. Absolutely. I'd send a 90-year-old grandmother on a road trip with Autopilot because once they get acclimatized to it, it's reproducible and it enhances their safety bubble. Turn on the Interstate at 70 miles an hour because it can."* (R084)

In contrast, 42 % of participants trusted FSD Beta, of which 18 % self-reported a lot of trust in FSD Beta.

Figures 1a and 1b provide a distribution of the responses obtained for the questions pertaining to the knowledge about the operation of Autopilot and FSD Beta (top) and trust in Autopilot and FSD Beta (bottom).

3.4. Factors of trust and perceived safety

The analysis of the interview data resulted in the extraction of main categories and subcategories. The categories represent the factors of trust and perceived safety. In total, we identified three main categories and twenty-nine subcategories, representing human-operator-related (12), automation-related (10), and environmental factors (7). Of the human-operator related factors, we identified one factor representing three personality traits (i.e., locus of control, neuroticism, sensation seeking), two factors capturing states of human operators (i.e., providing adequate supervisory control, situational awareness and workload), five cognitive factors (i.e., (age-

Table 1
Factors of trust and perceived safety, meaning, and seed terms / sources

Factors		Meaning	Seed terms / sources
Main category	Subcategory		
	Human traits		
	Personality		
	Locus of control	Two-dimensional personality trait, with internal ('Self'), implying to accept personal responsibility for traffic-related outcomes, and external ('Other drivers') implying to attribute responsibility of traffic-related outcomes to other participants and unforeseen events in environment	'I am responsible', 'driver is responsible', 'responsibility', 'control', 'fault' (Özkan & Lajunen, 2005), 'responsible', 'responsibilities', 'other drivers'
	Neuroticism	Perceived negative emotional states during manual driving, such as nervousness, tenseness, or stress	'always nervous', 'generally nervous', 'conscious aware', 'generally moody' (Lajunen, 2001), 'generally tense', 'generally stressed' (Matthews & Desmond, 2000), 'trouble', 'worrying', 'nerves', 'depressed', 'confrontation', 'tension', 'uneasy', 'awkward', 'pain'
	Sensation seeking	Tendency for sensation seeking, exploring limits of technology, pushing its boundaries, and engaging in risky or sensational driving experiences	'adrenaline', 'cliff', 'exciting' (Hoyle, Stephenson, Palmgreen, Lorch, & Donohew, 2002), 'racing', 'unsafe passing', 'speeding' (Arnett, Offer, & Fine, 1997), 'race car', 'risk taking', 'impaired driving', 'non-use of seatbelts', 'running red lights' (Zuckerman, 1994), 'obsession', 'youngster', 'sensations', 'fascination'
	Human states		
	Providing adequate supervisory control	Establishing perceived safety by adequate supervisory control (i.e., supervising automation)	'supervise', 'monitoring', 'monitor'
	Situational awareness and workload	Reduction in workload and increase in situational awareness as a result of using automation	'awareness', 'situational awareness', 'mental capacity', 'mental workload', 'demand' (Hart, 2006)
	Cognitive factors		
Human operator-related	(Age-related) impairments	Physical and visual impairments	'eyesight', 'my age', 'my eyes', 'visual issues', 'fatigued', 'intoxicated', 'tired', 'impaired', 'reflexes', 'impairments', 'dysfunction', 'impairment', 'handicaps'
	Experience (familiarity)	Learning to trust automation over time, experiencing it in various situations and operational design domains (ODDs)	'over time', used it before', 'experience', 'experienced', 'experiences', 'learning', 'it's Beta'
	Mental model	Understanding of capabilities and limitations of automation	'know how it works', 'capabilities', 'mental model', 'ODD', 'operational design domain', 'knowing where to trust', 'highway', 'city' 'capability', 'capable', 'operational', 'limitations', 'boundaries', 'restriction'; 'limitation', 'constraints', 'limiting', 'limits'
	Technical expertise	Technical expertise, contributing to understanding of software as technology being prone to failure, reducing likelihood of initial miscalibrated overtrust	'never trust software', 'software can always fail', 'software person', 'tech-savvy', 'computer background', 'engineering', 'tester', 'software', 'engineer'
	Trust	Trust in automation to perform driving task safely, contributing to perceived safety, distracted, mind-off and fatigued driving, more laxed placement of hands on steering wheel, and feet off pedals	'trust', 'distracted', 'rely on', 'trustworthy' (Körber, 2018), 'comfortable'
	Emotive factors		
	Perceived safety		
	Feeling safe	Positive valences associated with feeling safe	'safe', 'relaxed' (Xu et al., 2018), 'happy', 'relaxation', 'relaxing', 'comfortably', 'glad'
	Feeling unsafe	Negative valences associated with feeling unsafe	'nervous', 'anxious', 'anxiety', 'unsafe', 'stressed', 'risky' (Xu et al., 2018), 'dangerous', 'fearful', 'worried', 'frustrated', 'unhappy', 'impatient', 'hesitant', 'discomfort', 'nervousness', 'headache', 'fatigue', 'harmed', 'riskier'; 'danger', 'threat', 'threatening', 'angry', 'risk', 'terrifying', 'afraid', 'stressful'
	Capability		
	Confidence	Automation knowing what it does, with assertive (aggressive) behavior indicating high confidence, and hesitant and cautious behavior indicating low confidence	'knows what it does', 'confident', 'is afraid', 'confidence', 'assured'
Automation-related	Safety benefits	Safety benefits due to vehicle's assisted and automated driving functionalities, and associated 360-degree sensing capabilities, contributing to avoidance of blind-spot related, rear-end, side-swipe, and collisions involving other road users, objects, and events in environment detected by automation	'saved', 'safer than', 'second set of eyes', 'prevented', 'avoided', 'combination', 'protection', 'hazard', 'blind spot', 'rear-end', 'crash', 'crashes', 'collision', 'collisions', 'accident', 'accidents', 'parked', 'truck', 'trucks', 'trailer', 'tractor'

(*continued on next page*)

Table 1 (*continued*)

Factors		Meaning	Seed terms / sources
Main category	Subcategory		
	Longitudinal and lateral distance to other road users	Longitudinal and lateral vehicle distance, with smaller distances to objects and events associated with higher trust in automation	'dead-centered', 'distance', 'stopping distance', 'following distance', 'mid line'
	Natural and human-like driving behavior	Natural and human-like driving behavior	'human-like', 'drive like we do',
	Predictability	Correspondence between actual and expected automation performance and accuracy of automation, doing what it is supposed to do, meeting expectations of drivers	'predictable', 'predictability', 'accurate', 'accuracy' (Muir, 1994; Rempel et al., 1985), 'predictably', clarity', 'smoothness', 'smooth'
	Reliability	Reliability of automation by consistent performance proven over time	'reliable', 'reliability', 'earned', 'unreliable', 'unreliability', 'accurate', 'dependable', 'relying', 'consistent', 'consistency', 'proven', 'durability', 'effectiveness', 'robustness', 'robust', 'dependability'
	Features		
	Driver monitoring system (DMS)	Camera-based DMS monitoring participants' eye gaze behavior	'cabin camera', 'driver monitoring'
	Explainability	Explaining behavior and performance of automation in retrospective	'explain', 'opaque' (Endsley et al., 2003), 'reason', 'describe', 'why', 'explaining'
	Transparency	Transparency ('observability') about intentions, performance, future actions, and reasoning process of automation	'visibility', 'understandable', 'transparency', 'transparent' (Endsley et al., 2003)
	Physical environment		
	Software updates	Software updates changing previously functional behavior leading to regressions, or improvements of automation behavior, contributing to changes in mental model	'software update', 'software updates', 'updates', 'updated', 'updating', 'advance', 'improvements', 'improve', 'improvement', 'improving', 'improved', 'upgrade', 'upgrades', 'upgrading', 'regression', 'bug', 'bugs'
	Stability	Using automation in stable operational design domains (ODDs)	'stable', 'stability'
	Advertising, education, training		
	Crash statistics	Statistics supporting crash reduction benefits of Tesla vehicles equipped with partially automated driving functionality	'statistics', 'data, 'statistic'
Environment-related	Email	Email from Tesla informing drivers prior to first FSD Beta use about automation potentially doing the wrong thing at worst time	'doing the wrong thing at the worst time' (Nordhoff et al., 2023), 'email'
	Insufficient training	Insufficient training at dealerships on proper automation use, contributing to initial inadequate automation use	'lack of training', 'insufficient training', 'practice', training'
	In-vehicle warnings	In-vehicle warnings alerting drivers to be vigilant and supervise automation to be able to respond to take-over requests	'warning', 'warnings'
	Release notes	Release notes informing drivers about changes in automation behavior	'release note', 'release notes'
	Social influence	Overselling capabilities of Autopilot and FSD Beta via social media, contributing to unrealistic expectations about actual capabilities of automation	'social media', 'Youtube', 'videos'

related) impairments, experience (familiarity), mental model, technical expertise, trust), and one emotive factor (i.e., perceived safety). Of the automation-related factors, we identified six factors representing the automation's capability (i.e., confidence, safety benefits, longitudinal and lateral distance to other road users, natural and human-like driving behavior, predictability, reliability), and three factors capturing features of the automation (i.e., driver-monitoring system, explainability, transparency). Of the environmental factors, we identified two factors related to the physical environment (i.e., software updates, stability), and six factors related to advertising, education, and training (i.e., crash statistics, email, insufficient training, in-vehicle warnings, release notes, social influence).

Table 1 presents an overview of the factors influencing perceived safety and trust, their meaning, as well as the seed terms / sources used to examine their association with trust and perceived safety as presented in Table 5. A descriptive presentation of the results in Table 1 is provided in the sections 3.6–3.8.

3.4.1. Human operator-related

Human operator-related factors were divided into four main factors pertaining to human operator traits, states, cognitive, and emotive factors. Regarding personality traits, participants expressed a **preference to maintain control during manual driving** due to a general awareness of risks in traffic and a lack of trust in other drivers. They experienced **negative emotional states during driving**, such as nervousness, tenseness, or stress, which is indicative for a **neurotic personality trait.** We also found a **tendency of participants to seek sensations**, exploring the limits of the technology, pushing its boundaries, and engaging in risky or sensational

Table 2
Sample of quotes representing human operator-related factors of trust and perceived safety.

Main category	Subcategory	Sample quotes	
		Trust	Perceived safety
Human operator	Traits		
	Personality		
	Locus of control	*„You can't trust other people. You always assume that something bad can, will happen to you. I generally drive that way. I find it relatively dangerous. That's the attitude I have on driving."* (R016)	*„No matter when I'm driving, it's never safe. A lot of the things can go wrong so I do not feel safe, and that's because you're driving with other people that make mistakes. People make mistakes."* (R036)
	Neuroticism	*„I am a cautious person, so I never trust it completely."* (R100)	*„I probably never feel very safe in a car. That's me personally. I'm generally nervous about riding and driving. I'm much more conscious, aware, nervous. Most people don't think twice, and then people have the wrong ideas about what is and what isn't dangerous."* (R002)
	Sensation seeking	*„It's fun for me. It can be exciting. There's a highway 17. It's a legendary road for being twisting and accidents, and I tell people 'I let the car drive on South Peanut for me. Look at me.' I'm afraid, but can't complain.' Kind of brave."* (R042)	*„I love it. I love it. If it's driving me off a cliff, I could care less. That's what I pay for. I could care less. I'm an engineer. I love testing these types of systems, so I feel safe."* (R062)
	States		
	Providing adequate supervisory control		*„I feel safe because I can take over at any time."* (R011)
	Situational awareness and workload		*„Do you feel safe? 'Yes', and that is because it enables me to pay more attention to my surroundings."* (R041)
	Cognitive factors		
	(Age-related) impairments	*„I trust the computer features more than I trust myself. I have had two accidents where I never had any before. Maybe I'm getting a little older."* (R021)	*„I actually feel more safe now. At my age, my reflexes aren't as good as they were 10, 20 years ago. My vision, it's not as good as it used to be. I'm having hearing difficulties. I feel that my safety is improved with my Tesla on the road out here."* (R086)
	Experience (familiarity)	*„My trust has changed with experience for sure. I trust the system, the more I have had a chance to use it."* (R063)	*„Autopilot is very safe. I have closed 30,000 miles, and 99 % of the time I feel very safe."* (R062)
	Technical expertise	*„You probably need to rephrase the question 'Do I trust the system too much?' The question should be: 'How much do I understand AI?' How well do you know the technology yourself? What is your background? Are you familiar with such technologies?' I'm not an AI guru, but I have a reasonable understanding that I say 'OK. This technology masters certain situations. It's not this black box.'"* (R004)	*„Ultimately, it's software, and software always has problems. I would never trust software completely. There's always safety issues."* (R003)
	Mental model	*„I learned to understand it. All its strengths and the weaknesses. My trust is a lot more than before because when you first get Autopilot, it's not easy, but now we understand each other."* (R073)	*„I feel safe because I understand those boundaries. I understand the attention it required. You have to know the boundaries."* (R031)
	Emotive factors	Feeling unsafe	Feeling safe
	Perceived safety	*„Anxiety. It's when you're driving in the rain, and you can't see well. High-adrenaline. Anxiety."* (R007)	*„When I feel safe, it's perfectly fine. I feel bored where I have a lot more time to think about what's going on around me."* (R054).

driving experiences. The state of human operators was captured by two main factors pertaining to establishing safety by providing adequate supervisory control. Further, a **decrease in workload** and an **increase in situational awareness** was reported as a result of using Autopilot. We also identified several cognitive factors. P**articipants with (age-related) visual and physical impairments mentioned a higher trust and perceived safety using the automation in comparison to driving manually**. Participants' **technical expertise and understanding** of software being prone to failure emerged as another factor impacting perceived safety and trust. Most of the participants had a **correct mental model of Autopilot and FSD Beta** as partially automated driving systems that require permanent supervision by human drivers, understanding **what the automation can** and **specifically cannot do**. They **learned to trust the automation over time**, **experiencing** it in various situations and operational design domains (ODDs). Participants mentioned **positive and negative valences of perceived safety**. Feeling safe was associated with positive emotional states, such as feeling relaxed, comfortable, happy, held, bored, or energized. Feeling unsafe was associated with negative emotional states, such as feeling stressed, anxious, nervous, or being cautious. Table 2 presents illustrative quotes representing the human operator-related factors of trust and perceived safety.

3.4.2. Automation-related factors

We identified automation-related factors representing the automation's capability and features. The **hesitant and cautious behavior of the automation** was indicative for a **lack of confidence of the automation**. Participants mentioned the **safety benefits of using the automation** due to the **automation augmenting the skills of human drivers**, and **avoiding crashes**, such as **blind-spot, rear-end**, and **side-swipe collisions**, as well as **detecting vulnerable road users (VRUs)**, particularly during **distracted driving due to intoxication or fatigue**. Another subcategory that emerged from the data analysis was the **'longitudinal and lateral**

Table 3
Sample quotes representing automation-related factors of trust and perceived safety.

Main category	Subcategory	Sample quotes	
		Trust	Perceived safety
Automation-related factors	Capability		
	Confidence	*„Where it's not sure what to do. It is hesitant, and it could cause an accident because people are behind. It's slowing down, and you'll feel it go and stop. You don't want 100 % trusted."* (R008)	*„My big thing about Beta is its lack of confidence when performing maneuvers. It's lack of confidence makes you feel unsafe."* (R071)
	Safety benefits		
	Automation augmenting skills of drivers	*„I trust it now. If you come back home very late at night and you doze off, it's really nice to have Autopilot, probably because it does save you. I do see other cars. They might be drunk because they're swerving. Autopilot probably prevented a crash."* (R033)	*„A second set of eyes makes you feel safer."* (R079)
	Blind-spot deleted crashes		*„When I was driving, my car started alerting me, beeping and there's a red icon flashing on the screen telling me to take control immediately, and then I looked in my blind spot: A car in the next lane veered into my lane. I didn't see it, but Autopilot saw it, and reacted. It shifted my car away. I was very impressed."* (R091)
	Detection of vulnerable road users (VRUs)		*„It was dark, and I turned a corner and suddenly a guy walked in front. I would have hit him. I didn't see him until he was looking right into the windshield laughing."* (R093)
	Rear-end collisions		*„I was on a road trip down to Mississippi. The car would not let me make the lane change, had a really stiff force on the wheel. I couldn't really turn it. I saw that it sees the car. He would have either rear-ended me. The technology saved me from a minor Fender Bender."* (R087)
	Side-side collisions	*„It's watching out for things, and it reacts when people pull in your lanes, or do something silly. Autopilot was able to either slow down or in some way move out of the way."* (R061)	*„It has saved me from being sideswiped by a car merging."* (R093)
	Longitudinal and lateral distance to other road users	*„I usually like about 3, 5 cars in front and following. Good distance, and I have mine set to follow 5 or 6 cars because I don't trust it."* (R008)	*„It needs to stay away from trucks. It needs to figure out what it's doing with the stopping distance from the cars in front of me because it feels close to unsafe."* (R070)
	Natural and human-like driving behavior	*„So improving the trust. It has to turn more into a natural driver concept."* (R042)	*„Putting in some more natural or human feeling would make me feel safer."* (R050)
	Reliability	*„Super consistent. Lots of trust built up over time. It's so good at what it does. That is the level of trust it has gained, rightly or wrongly. It has induced that level of trust by its consistency."* (R084)	*„I 100 % feel safe. It's so solid. I even did a road trip. It was seven hours of driving to Northern California, and then seven hours of driving to Southern California, and it's perfect as it can be."* (R055)
	Predictability	*„Most of that trust and feeling of safety is wrapped up in its predictability, and so I know what I am expecting the car to do, and when the car does it, then I'm feeling safe."* (R037)	*„Do I feel safe? No, because it's still too unpredictable. That would be the short answer."* (R081)

(*continued on next page*)

Table 3 (*continued*)

Main category	Subcategory	Sample quotes	
		Trust	Perceived safety
	Features		
	Driver monitoring system (DMS)	*„Some of the risks are that you can get a little bit complacent. The cabin camera helps to avoid some of that because if the driver isn't paying attention, it triggers an alert. I've written an entire email while driving in the Model X where I know the camera is not watching me. I can't do that in the Model 3 because the camera is watching me."* (R054)	
	Explainability	*„If you could see the decisions made, and say 'Accident or object avoidance'. If we could see why it makes the choices that it does."* (R040)	*„I would definitely feel safer if I understood why the car was making certain decisions. A little bit more transparency there would be good. Sometimes it feels like an opaque black box."* (R053)
	Transparency	*„I would trust version 10 more than version 9 just because it shows so much more on the screen."* (R043)	*„I feel safer because I see the perception on the screen, and that gives me confidence of knowing exactly what it is seeing compared to blindly believing."* (R087)

distance' between the vehicle and objects and events in the environment. Another subcategory was the lack of **'human-like and natural automation behavior**'. The **reliability** and **predictability** was associated with the consistency and accuracy of the automation's behavior. Regarding the features of the automation, we found that the **camera-based driver monitoring system** tracking participants' gaze behavior **reduced complacency**, and **encouraged safer use.** The **in-vehicle display** supported the **explainability of the automation behavior**, and provided **transparent information** about its **intentions, performance, future actions,** and **reasoning process**. Participants considered the **visualizations on the screen visually demanding**, **distracting from** the task of **driving**. Sample quotes representing the automation-related factors of trust and perceived safety are presented in Table 3**.**

3.4.3. Environmental factors

Environmental factors are represented by the physical environment, and by advertising, education, and training. Regarding the physical environment, participants mentioned the **use of the automation in stable operational design domains (ODDs)**. **Software updates** resulted in **software regressions or bugs leading to dysfunctional behavior**, or to **improvements** in automation behavior. We identified several sub-categories representing advertising, education, and training. **Crash statistics enhanced** the **perceived safety benefits** of the vehicles equipped with Tesla's partially automated driving systems. **Informing** participants prior to the first use **via email** about FSD Beta potentially **'doing the wrong thing at the worst time' heightened participants' awareness to monitor the automation, avoiding complacency**. **Insufficient training at dealerships** resulted in **initial inadequate use** of the automation. **Information** that participants received **via the vehicle** (e.g., **warnings**) **informed them accurately** about their role and responsibilities. The **name of Autopilot and FSD Beta** was **misleading** for some participants, contributing to initial overtrust and low perceived safety. **Release notes** informing participants about changes in the behavior of the automation by software updates **improved the understanding of system** capabilities and limitations. **Watching others using the automation contributed to building more accurate expectations about the behavior of the automation**. Sample quotes representing the environmental factors of trust and perceived safety are presented in Table 4.

Table 4
Sample quotes representing environmental factors of trust and perceived safety.

Main category	Subcategory	Sample quotes	
		Trust	Perceived safety
Environmental factors	Physical environment		
	Software updates	*„I started to trust the technology more with each update. When they start doing unprotected left turns, I would clap in my car like 'You did a great job. That was perfect.'"* (R091)	*„When they introduce something new, something else breaks. It can be safe in certain ways, but at the same time it creates anxiety and performs maneuvers that makes me feel unsafe."* (R071)
	Stability	*„I trust FSD in many ways, but it's still so bad. I wouldn't want to put it in a stressful situation. San Francisco would be crazy."* (R042)	*„It depends on the circumstance. If it's just open highway with very little traffic, obviously it feels safe."* (R016)
	Advertising, education, training		
	Crash statistics	*„I'm going to have to rely on Tesla's public numbers. I'll be like, 'OK, if those are the numbers, I'll trust that.'"* (R007)	*„I feel safe, but it also has to do with good test results, with security tests."* (R039)
	Email	*„By definition, I'm not allowed to trust it. They say 'Keep your hands on the wheel. Pay attention at all times.' The car may throw itself off a bridge at any moment."* (R017)	*„I would say my perceived safety is very low. It definitely hasn't gone up much. I still expected it to do all the wrong things at all the wrong times."* (R045)
	Insufficient training	*„When I first used it, I didn't have any hands on the wheel because I didn't know how to use it, and then it yelled at me to*	*„When I started using Autopilot, there was very little training on how to use Autopilot. I was making many, many mistakes. I*

(continued on next page)

Table 4 (*continued*)

Main category	Subcategory	Sample quotes: Trust	Sample quotes: Perceived safety
		put the hands on the wheel, but I didn't know how to use it." (R052)	*didn't really understand the functionality. It used to be a scary drive initially because of one simple reason: I was not trained on what a following distance is. It was set as a default of 1. That used to scare me completely. Once I changed the following distance to 6, using Autopilot was absolutely the best thing possible for me."* (R035)
	In-vehicle information	*„I trust these systems. Everything Tesla sets, everything it says in the documentation in the vehicle when you enable these features."* (R005)	*„It is Beta software and they give us all the warnings about how it could be dangerous, how that car might do. They say red on the screen, every update you may have, you must be prepared because the car may do the worst thing at the worst time."* (R043)
	Release notes	*„They started giving release notes. It helps build the level of trust."* (R061)	*„Every time I read the release notes, I am able to understand what they improved, and reading the long list of improvements has a high impact on my perception of the safety."* (R037)
	Social influence	*„I don't trust Full Self-Driving, especially because when my husband's tried it, and then said 'Ohh, it's really bad in this area', and 'Ohh, it almost killed me in this area.'"* (R059)	*„I'm not willing to do it until I hear other people are doing it successfully. That's my criteria. I follow a lot of other FSD Beta testers who put out videos and they're still struggling."* (R100)

Table 5
Association with trust and perceived safety.

Factors		Co-occurrence, *n*						Occurrence, *n*		
		Trust		Delta	Perceived safety		Delta	Trust	Perceived safety	Delta
Main category	Sub-category	Low	High	High-low	Low	High	High-low			Perceived safety − Trust
Human operator-related	Traits									
	(Age-related) impairments	0	16	**16**	31	30	**−1**	16	61	**45**
	Locus of control	8	62	**54**	90	129	**39**	70	219	**149**
	Neuroticism	3	18	**15**	32	40	**8**	21	72	**40**
	Sensation seeking	1	3	**2**	45	14	**−31**	4	59	**55**
	States									
	Providing adequate supervisory control	1	31	**30**	55	120	**65**	32	172	**120**
	Situational awareness and workload	0	9	**9**	18	26	**8**	9	45	**36**
	Cognitive factors									
	Technical expertise	4	48	**44**	56	74	**18**	52	130	**78**
	Experience (familiarity)	2	95	**93**	133	134	**1**	97	267	**170**
	Mental model	10	170	**160**	184	197	**13**	180	381	**201**
Automation-related	Capability									
	Confidence	3	70	**67**	44	63	**19**	73	107	**34**
	Safety benefits	6	114	**108**	232	232	**0**	120	464	**344**
	Longitudinal and lateral distance to other road users	3	26	**23**	43	78	**35**	29	121	**92**
	Natural and human-like driving behavior	1	9	**8**	7	13	**6**	10	20	**10**
	Reliability	4	53	**49**	48	59	**11**	57	107	**50**
	Predictability	2	32	**30**	68	46	**–22**	34	114	**80**
	Features									
	Driver monitoring system (DMS)	0	0	**0**	4	4	**0**	0	8	**8**
	Explainability	7	79	**72**	180	132	**−48**	86	312	**226**
	Transparency	0	9	**9**	5	13	**8**	9	18	**13**
Environmental	Physical environment									
	Stability	0	3	**3**	7	5	**−2**	3	12	**9**
	Software updates	5	167	**162**	102	180	**78**	172	282	**110**
	Advertising, education, training									
	Crash statistics	4	55	**51**	72	78	**6**	59	150	**91**
	Email	0	6	**6**	14	24	**0**	6	38	**24**
	Insufficient training	0	13	**13**	19	25	**6**	13	44	**31**
	In-vehicle warnings	2	13	**11**	19	25	**6**	15	44	**25**
	Release notes	0	8	**8**	0	1	**1**	8	1	**7**
	Social influence	3	18	**15**	23	29	**6**	21	52	**29**
Total *n*		69	1127	1058	1531	1771	230	1196	3300	2077

3.5. Association between factors

The association between the factors of trust and perceived safety was examined by counting the co-occurrence of seeds representing the factors, and trust and perceived safety, respectively.

The largest difference in the total number of mentions of trust and perceived safety was obtained for the subcategory ‘safety benefits’ ($n = 344$) A larger difference is associated with a stronger association. The total number of the seeds representing perceived safety and ‘safety benefits’ was found 464 times, while the combinations of the seed terms representing trust and ‘safety benefits’ was found 120 times. The second-largest difference was obtained for the subcategory ‘explainability’ ($n = 226$). The combination of the seeds representing ‘explainability’ and perceived safety was found 312 times, while the seeds for trust and explainability appeared 86 times.

The smallest differences were obtained for the subcategory ‘release notes’ ($n = 7$), with the combination of the seeds for trust and release notes being found 8 times in comparison to the seeds for perceived safety and explainability being observed once. The second-smallest difference was found for the sub-category ‘driver monitoring system’ ($n = 8$): The combination of the seeds for perceived safety and driver monitoring was found eight times, while we did not find the combination for trust and driver monitoring.

The largest differences between low and high trust were found for the subcategory ‘software updates’, with a total difference equaling ($n = 162$), with the seed terms representing low trust being mentioned 5 times, and the seed terms representing high trust 167 times. The second largest difference was found for the subcategory ‘mental model’ ($n = 160$), with the seeds representing low trust being mentioned 10 times, and the seeds representing high trust being mentioned 170 times.

3.6. Relationships between perceived safety, trust, and automation use

We also examined the self-reported relationships between perceived safety, trust, and automation use. Participants **felt safer because** they **trusted the automation. Trust and perceived safety promoted eyes-off road, hands-off-wheel, and feet-off-pedal driving**, and the **self-reported frequency of automation use.** Table 6 provides an overview of illustrative quotes representing the relationship between trust, perceived safety and automation use.

Table 6
Sample quotes representing relationship between trust, perceived safety, and automation use.

Variable	Dependent variable	Automation use				
Independent variable	Perceived safety	Secondary task engagement	Complacency	Removing hands from steering wheel	Removing feet from pedals	Frequency of use
Trust	*„I feel safe enough that I can take my eyes off the road for a couple minutes and I would not do that if I did not trust it, if I did not feel safe."* (R055)	*„I keep some antacids in my car. I've got heartburn, and I can open it with two hands very comfortably. I can do that comfortably, and I can trust the car to drive."* (R056)	*„How willing you are to trust the car drives backs into how much attention you pay to the system. If you are very comfortable with the driving agent, you can set more of your attention to other vehicles, and lane lines rather than the actual driving software."* (R002)	*„If I start not having to hold the wheel with both hands, that means I've seen it operate enough that my guard gets let down a little bit and I trust him more."* (R031)	*„Over time it grew more confidence with the car that I didn't need to have my foot close to the brake. If I need to hit the brake usually in the emergency, the Tesla will already be slowing down or braking for me. Even if my foot is not right closer to brake, I have enough time to engage the brake fully."* (R073)	*„I have a lot of trust. I use it all the time. I would hate to not have it every day."* (R003)
Perceived safety		*„Feeling safe? It makes you pay a lower amount of attention. I am more relaxed, and so I'm willing to look out the side window for a few seconds, looking at the scenery."* (R096)	*„I feel safe enough to become complacent. Autopilot is so safe on the highway."* (R074)	*„I actually take my hands off the wheel because I feel safer doing that. I actually feel better taking my hands off the wheel and then look away so I don't accidentally steer."* (R044)		*„I don't use it in situations where I feel unsafe. Those are very rare, so I use it all the time."* (R067)

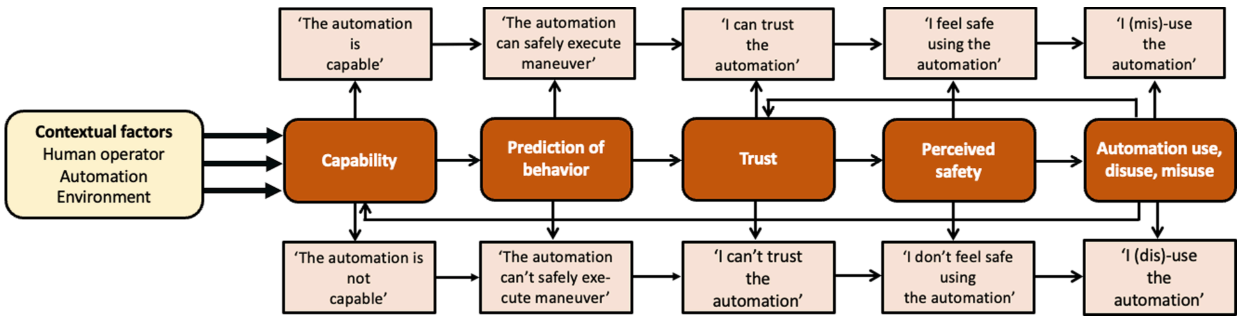

Fig. 2. Conceptual, process-oriented model, PTS-a, for the relationships between contextual factors, trust, perceived safety, and automation use.

3.7. *A synthesis*

The results of the data analysis informed the development of a novel conceptual, process-oriented model, titled PTS-a, which stands for predicting trust in and perceived safety of automation use, as presented in Fig. 2. The model identifies the human-operator-, automation-related, and environmental factors influencing a human operator's assessment of the automation's capability. The human operator's capability assessment influences the prediction of the behavior of the automation. If human operators decide that 'the automation is capable', they are more likely to believe that 'the automation can execute the maneuver safely'. In the framework, this is the basis for the development of trust in the automation. If 'the automation can execute the maneuver safely', human operators develop trust in the automation – 'I can trust the automation'. Trust forms the basis for the development of perceived safety 'I feel safe using the automation'. This in turn determines automation use, including misuse, and disuse. The model also hypothesizes a direct effect of trust on automation use. Furthermore, it includes a feedback loop, with automation use initiating the human operator's reasoning process with regards to whether the automation can be regarded as capable.

4. Discussion

This study reports the results of an interview study conducted with 103 users of real-world partially automated driving systems to examine the factors of trust and perceived safety, and the relationship with automation use. Based on their co-occurrences represented by specific seeds for trust and perceived safety, we also explored the association between the factors of trust and perceived safety. Finally, we synthesized the results of this study, and propose a conceptual model for the relationships between contexual factors, trust, perceived safety, and automation use.

4.1. *What are human operator-related, automation-related-, and environmental factors of trust and perceived safety?*

In total, our analysis resulted in the identification of three main categories capturing twenty-nine human operator-related, automation-related, and environmental factors of perceived safety and trust. Overall, our automated count analysis has shown that all factors are more strongly associated with perceived safety than with trust in the automation.

4.1.1. *Human operator-related factors*

We identified four main human operater-related factors represented by traits, states, cognitive, and emotive factors. We found that human operator's personality was associated with trust and perceived safety. While studies have found that neurotic people were more likely to have (fatal) accidents (Eysenck, 1965; Lynn & Hampson, 1975; Özkan & Lajunen, 2007), our participants with neurotic behavior during manual driving reported lower trust and perceived safety when the partial driving automation was engaged. Neuroticism has been found to have a negative effect on trust in automation (Zhang et al., 2020). We found a positive association between locus of control and trust and perceived safety, with locus of control being associated with a higher number of seeds representing higher than lower trust, and higher than lower perceived safety. At the same time, the combination of the seeds representing distrust (or low trust) and locus of control received the second-highest hits ($n = 8$) among all combinations representing distrust. Future research should validate the relationship between neurotic behavior during manual driving and trust and perceived safety using partial driving automation, and identify further laymen-friendly seed terms to identify psychological constructs being pivotal for automation use in unstructured text data.

Participants' technical expertise contributed to an understanding of software as technology being prone to failure, initially mitigating the likelihood of overtrust in the automation. Participants had a correct understanding of the systems requiring human supervision and intervention, which differs from the study of Mueller et al. (2023) in which users of partially automated driving did not have an accurate understanding of the limits of the system.

However, a correct mental model of the system did not prevent complacency and misuse. Trust and perceived safety promoted the likelihood of complacency, secondary task engagement, and hands-free driving. Given that one of our participants mentioned that Autopilot "has induced that level of trust by its consistency", our data may support the notion of automation-induced complacency. The notion of automation-induced complacency was first introduced in automated aircraft, with the pilot and crew failing to monitor the automation in highly reliable automated environments (Parasuraman, Molloy, & Singh, 1993). Secondary task engagement can lead to a crash or other critical incidents if drivers miss a warning or message from the automation, failing to take over control in time (Campbell et al., 2018). Informed by a review of literature investigating vigilance in the context of partial driving automation, McWilliams and Ward (2021) concluded that "there will be cognitive underload and subsequent vigilance decrements as auto manufacturers continue to produce vehicles capable of PAD. What is less clear is the underlying mechanism" (p. 9), and its impact on traffic safety. Future research should be conducted to gain a better understanding of why, when, where, and with whom this phenomenon occurs, and how drivers can be supported in using the system more properly.

Participants' mental model developed from experiencing the automation in various situations and environments. It was more strongly associated with higher than lower trust. This is in line with studies showing a positive impact of user's mental model on trust (Kazi, Stanton, Walker, & Young, 2007), but contrasts other studies, which in trust influenced a user's mental model (Rudin-Brown & Noy, 2002). Experiencing the automation had a positive impact on trust and perceived safety. Experiencing the automation was more strongly associated with high than low trust, which may indicate that experience impacts trust. Participants experiencing partial driving automation in controlled test track studies had a high level of trust before, during, and after experiencing the automation (Blanco et al., 2015). Conversely, experience contributed to a small difference between low versus high perceived safety. This finding implies that experience may not significantly improve perceived safety, which may be counterintuitive. In Russell et al. (2021), experience increased perceived safety of using a lateral driving feature. We recommend future studies to further explore the association between experience and perceived safety.

Participants reported a high level of self-reported trust in Autopilot, with 55 % of participants stating to trust Autopilot a lot, supporting research from Lenneman, Mangus, Jenness, and Petraglia (2020) who have shown a high self-reported trust in Driver Assistance Technologies (DAS) among experts. In the conceptual framework by Ekman et al. (2017), trust in the pre-use phase is not appropriately calibrated (i.e., is either too high or too low), but calibrates in accordance with the automation's actual competence and is fairly stable over time until the occurrence of an incidence that may result in a decrease in trust. The framework does not consider an inappropriate trust calibration in the performance phase, which is based on the performance of the automation, capturing the long-term perspective. Given that some of our participants were experienced with the automation, reported high levels of trust, and misused the automation, future research should discuss the likelihood of mis-calibrated trust (over-trust) in the performance phase.

Another interesting finding of our study addresses the perceived safety of elderly and impaired participants who reported a higher perceived safety with the automation engaged in comparison to driving manually. The acceptance of elderly people was high if the automation demonstrates high reliability (Haghzare, Campos, Bak, & Mihailidis, 2021). In Haghzare et al. (2022), participants with dementia rated the perceived safety of fully automated vehicles higher than their own safety as drivers. We recommend future research to conduct naturalistic, on-road driving and interview studies with elderly and impaired populations to understand how impairments affect trust and perceived safety as well as take-over performance in different, safety-critical conditions.

4.1.2. Automation-related factors

Automation-related factors were represented by the capability and features of the automation. Among the capability-based factors of the automation were the safety benefits that participants associated with the automation. Participants reported incidences of the automation "saving them", preventing different types of crashes. The actual crash reduction of vehicles equipped with partially automated driving functions is still largely unknown (Mueller et al., 2021; Noy, Shinar, & Horrey, 2018). Goodall (2023) found that the crash reduction rate of partially automated vehicles improved by only 10 % in comparison to cars equipped with assisted automated driving functions after controlling for differences in freeway driving.

We found a relatively small number of hits for the combination of seeds representing reliability, predictability, and trust and perceived safety, respectively. This is an interesting finding given that reliability and predictability were key factors influencing trust in the automation (Muir, 1994; Waung, McAuslan, & Lakshmanan, 2021).The reliability of the automation increased complacency, with participants monitoring the automation less as intended by designers (Banks et al., 2018; Parasuraman et al., 1993).

Increasing transparency and explainability of the automation's capabilities and process may positively impact trust and perceived safety. We found a lower number of counts for high than low perceived safety in combination with seeds representing the sub-category explainability, improving the explainability can lead to a higher perceived safety during partially automated driving. In line with Bhaskara, Skinner, and Loft (2020), the provision of visual information on the in-vehicle interface is visually demanding for some participants, negatively affecting their situational awareness and workload.

4.1.3. Environmental factors

The environmental factors were grouped into the physical environment, and advertising, education, and training. Overall, the automated count analysis has shown a smaller number of mentions of the seeds representing environmental factors in combination with perceived safety and trust. Advertising the safety benefits of partial driving automation via crash statistics promoted participants' perceived safety when using partial driving automation. Watching online videos contributed to both realistic and unrealistic expectations. Participants used online videos to support the development of their mental model about the capabilities and limitations of this technology. Videos contributed to trust calibration if they showed other users testing or experiencing the limitations of this technology. However, unrealistic expectations set through these online videos resulted in initial overtrust and inappropriate use of the automation. Furthermore, the email from Tesla warning drivers about FSD Beta potentially doing the wrong thing at the worst time appeared to have led to some sort of calibrated trust in the automation, and initial appropriate use of the automation. Future research is needed to investigate how sustained the effect is. In Beggiato and Krems (2013), participants receiving the most critical information about

Adaptive Cruise Control (ACC) directly after reading the system description had the lowest level of trust in the system. In Jieun Lee, Abe, Sato, and Itoh (2021), trust ratings did not differ substantially between the groups varying in the level of detail they received about the partially automated driving system. Consumer education can strengthen the relationship between participants' understanding of Advanced Driver Assistance Systems (ADAS) and confidence in the understanding of ADAS (Mason et al., 2023). A lack of training led to miscalibrated trust and inappropriate first use of the automation. In Russell et al. (2021), drivers who received training exaggerating the capabilities of the automation were more likely to take their hands off the wheel and their eyes off the road, but the effect of training was not significant when using the partial driving automation on the test track. In their study, training did neither influence perceived safety nor trust in a lateral driving feature. More research is needed to investigate the effect of information, training, and education on trust calibration, and use of partial driving automation.

4.2. How does trust in automation influence perceived safety, and automation use?

A conceptual, process-oriented model synthesizes the results of the data analysis, which posits that human operator-related, automation-related, and environmental factors as contextual factors influence trust, perceived safety and automation use. It provides insights into the cognitive-affective processes underlying automation use, treating trust and perceived safety as distinct cognitive and affective constructs.

In this model, contextual factors determine a human operator's assessment of the capability of the automation to handle a situation safely. The capability of the automation – 'The automation is capable' – leads to the formation of trust in the automation – 'I can trust the automation'. Automation confidence, competence, performance, or the expectation that the automation is working properly is central to building trust in the automation (Campbell et al., 2018; John D Lee & Katrina A See, 2004; Mayer et al., 1995; Muir, 1994; Waung et al., 2021). To help human operators correctly assess the capability of the automation, the automation may indicate its own assessment of its capability, e.g., using a capability indicator. The idea that the automation can assess its own competence is not new (Paardekooper et al., 2021).

The model treats trust as a cognitive belief that human operators form or develop as a result of logical reasoning and rational thinking of whether the automation can be trusted. Perceived safety, on the other hand, represents an emotion or feeling, which develops as a function of trust. The "cognitive emotion theory" or "cognition-leads-to-emotions" approach posits that beliefs are major determinants of emotions (Frijda, Manstead, & Bem, 2000; Lazarus, 1991a, 1991b). Other scholars have also referred to trust as affective construct capturing one's own feelings, instincts, and intuition (Kohn et al., 2021; Lewis & Weigert, 1985).

The model assumes that perceived safety – 'I feel safe using the automation' – can lead to (mis-)use of the automation. Our model hypothesizes that feeling unsafe can lead to (dis-)use of the automation. Both feeling safe and unsafe using the automation can lead to using the automation as intended by system designers, i.e., monitoring it adequately to be able to intervene in time when requested by the automation. In the behavioral adaptation model by Weller and Schlag (2004) behavioral adaptation (e.g., misuse, disuse) will occur if there is an objective and subjective enhancement of safety and a subjective enhanced utility of performing the behavior triggered by changes in the vehicle or the environment. Our model also includes a feedback loop, with automation misuse, disuse, and use influencing trust in line with Blanco et al. (2015), and initializing the process of assessing the capability of the automation.

4.3. Limitations and implications for future research

First, the data represents the subjective perceptions of participants. Experimental simulation studies, and studies using naturalistic on-road data should be conducted to further unravel the temporal nature of the relationship between trust, perceived safety, and use of partial driving automation.

Second, our sample is not balanced in terms of age, gender, income, and race / ethnicity. This bias must be corrected by future research as partially automated vehicles will be adopted by other parts of the population beyond the typical early adopters.

Third, we did not compute inter-coder reliability, which serves as important metric for the objectivity and reliability of the results of qualitative research.

Fourth, our research did not investigate the impact of other important factors of trust and perceived safety, such as demographics (e.g., age, gender, ethnicity), or trust propensity (Schaefer et al., 2014). We recommend future research to perform quantitative survey studies, and test the effect of these additional factors on perceived safety, trust, and use of partial driving automation.

Fifth, we did not validate the conceptual model. Future studies should validate it, critically assessing to what extent the conceptual model makes sense to scholars and practitioners from other disciplines, and proposing adjustments if applicable. This revised model can be presented at conferences and seminars (see Jabareen, 2009) until developments in the field trigger a new discussion that may lead to further changes in the model structure.

Sixth, it is plausible that some seeds may be inappropriate (Watanabe & Zhou, 2022), not fully capturing their subcategories due to the difficulty to measure psychological constructs using laymen-friendly language in qualitative studies. We recommend future research to further perform a seed term analysis of relevant constructs determining trust, perceived safety, and automation use.

Seventh, it is plausible that when participants were asked to subjectively describe and rate their trust and perceived safety, they may treat trust and perceived safety interchangeably. Future research should identify cognitive-related thoughts and beliefs pertaining to trust in automation and perceived safety to contribute to the operationalization of these constructs.

4.4. Conclusions

The study presents the results of semi-structured interviews with 103 users of real-world partially automated driving systems (i.e., Tesla Autopilot and FSD Beta), exploring the factors impacting trust in automation and perceived safety, and their relationship with automation use. We identified several human operator related, automation-related, and environmental factors influencing trust and perceived safety. An automated count analysis of the seeds was conducted to identify the factors and their association with trust and perceived safety. The factors that we identified were more strongly associated with perceived safety than with trust in partial driving automation. A correct mental model of the automation's capabilities and limitations can still lead to misuse of the automation, with participants failing to monitor the automation. A process-oriented model synthesizes the results of the analysis, representing the relationship between trust, perceived safety, and automation use. We recommend future research to define the thresholds of trust and perceived safety, which are conducive to misuse and disuse of partial driving automation.

Term	Definition	Author contributions
Conceptualization	Ideas; formulation or evolution of overarching research goals and aims	SN
Methodology	Development or design of methodology; creation of models	SN
Software	Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components	SN
Validation	Verification, whether as a part of the activity or separate, of the overall replication/ reproducibility of results/ experiments and other research outputs	SN
Formal analysis	Application of statistical, mathematical, computational, or other formal techniques to analyze or synthesize study data	SN
Investigation	Conducting a research and investigation process, specifically performing the experiments, or data/evidence collection	SN
Resources	Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools	SN
Writing – Original Draft	Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation)	SN
Writing – Review & Editing	Preparation, creation and/or presentation of the published work by those from the original research group, specifically critical review, commentary or revision – including pre-or postpublication stages	MH, SN
Visualization	Preparation, creation and/or presentation of the published work, specifically visualization/ data presentation	SN
Supervision	Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team	MH, SN
Project administration	Management and coordination responsibility for the research activity planning and execution	SN
Funding acquisition	Acquisition of the financial support for the project leading to this publication	SN

CRediT authorship contribution statement

Sina Nordhoff: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Marjan Hagenzieker:** Writing – review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to thank Lucas Suryana, PhD Candidate, Transport & Planning, Delft University of Technology, for his invaluable contributions in developing and discussing the Python code for the automated count analysis.

Appendix

Table A1
Questionnaire.

Question number	Question
Q1	Do you have the Full Self-Driving Beta (FSD Beta) feature? (1 = Yes, 2 = No)
Q2	Before the first time of using Autopilot and FSD Beta, did you watch / read / listen to information on how to use it? (1 = Yes, 2 = No)
Q3	Please mention the type of information you consulted on how to use Autopilot and FSD Beta (website of Tesla (www.tesla.com), car dealer / sales point, online communities and forums, YouTube videos, newspapers and magazines, friends, family, colleagues, driver manual)
Q4	Please describe your experience with using Autopilot and FSD Beta and the benefits and risks associated with using it. Please explain your answer.
Q5	Have your expectations of using Autopilot and FSD Beta been fulfilled? Why / why not?
Q6	Why do you use Autopilot and FSD Beta?
Q7	Did you ever stop using Autopilot and FSD Beta (for prolonged periods of time)?
Next, we would like to explore your perceptions regarding four general statements about the operation of Autopilot and FSD Beta.	
Q8	The current Autopilot does make driving autonomous. Is that correct? (1 = Yes, 2 = No, 3 = I don't know)
Q9	There are no safety issues with Autopilot. Is that correct? (1 = Yes, 2 = No, 3 = I don't know)
Q10	Autopilot is a hands-on feature. Is that correct? (1 = Yes, 2 = No, 3 = I don't know)
Q11	Tesla FSD Beta is safer than a human. Is that correct? (1 = Yes, 2 = No, 3 = I don't know)
With the next section, we would like to explore your perceptions of safety while using Autopilot and FSD Beta.	
Q12	Do you feel safe when Autopilot and FSD Beta is active? Why / why not?
Q13	What / how do you feel when you feel safe / unsafe? Please explain.
Q14	What is it about Autopilot and FSD Beta that is safe / unsafe? Please explain.
Q15	Now please remember the situation / s in which you typically feel unsafe when Autopilot and FSD Beta is active and describe these situations.
Q16	What can Autopilot and FSD Beta do to support your safety in Autopilot and FSD Beta? Please explain.
Q17	Does feeling safe / feeling unsafe impact how you use Autopilot and FSD Beta on your next drives / in the future? Please explain.
Q18	Has your perceived safety changed over time? If so, how?
With the next section, we would like to explore your trust in Autopilot and FSD Beta.	
Q19	How would you position your level of trust in Autopilot and FSD Beta. (1 = I don't trust it at all, 2 = I don't trust it, 3 = I neither don't trust it at all nor trust it a lot, 4 = I trust it, 5 = I trust it a lot)
Q20	What can Autopilot and FSD Beta do to support your trust in Autopilot and FSD Beta?
Q21	Does your trust / distrust in Autopilot and FSD Beta impact how you use Autopilot and FSD Beta on your next drives / in the future? Please explain.
Q22	Has your trust changed over time? If so, how?
Q23	When you do compare yourself with other participants, Autopilot, and FSD Beta, do you think you are … (1 = A much worse driver, 2 = A worse driver, 3 = Not a better nor a worse driver, 4 = A better driver, 5 = A much better driver)
With the next section, we would like to explore how you typically use Autopilot and FSD Beta.	
Q24	How do you typically place your hands on the steering wheel when Autopilot and FSD Beta is active? Please select the image that serves as best representation of your placement of your hands on the steering wheel when Autopilot / FSD Beta is active and explain your answer.

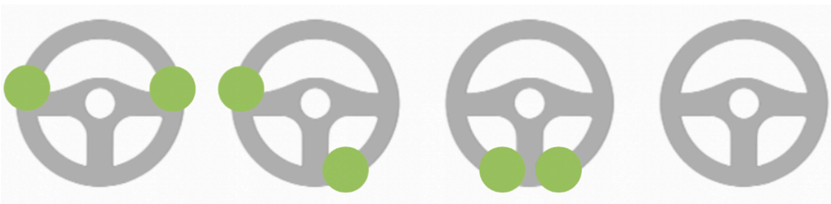

Figure is from Morando et al. (2021)

Question number	Question
Q25	Do you typically keep your hands on the steering wheel at all times?
Q26	Are you typically fully attentive and alert at all times?
Q27	How often do you typically engage in other secondary activities while Autopilot and FSD Beta is active? (Never, rarely, occasionally, frequently, always; monitoring the road ahead, talking to fellow travelers, observing the landscape, using the phone for music selection, using the phone for navigation, using the phone for calls, eating and drinking, using the phone for texting, watching videos / TV shows, sleeping)
Q28	Do you disengage Autopilot and FSD Beta? Why / why not?
Q29	Does Autopilot and FSD Beta disengage? When / in which situations?
Q30	How do you typically place your eyes when Autopilot and FSD Beta is active?
Q31	Do you typically keep your eyes on the road at all times?
Q32	Do you typically monitor the vehicle and its surroundings at all times?
Q33	How do you typically place your feet when Autopilot and FSD Beta is active?
Q34	Do you typically stay prepared to take corrective actions at all times?
Q35	Has your use of Autopilot (in terms of how you placed your hands on the steering wheel, eyes on the road, and feet) changed over time? If so, how?

References

Akalin, N., Kristoffersson, A., & Loutfi, A. (2022). Do you feel safe with your robot? Factors influencing perceived safety in human-robot interaction based on subjective and objective measures. *International Journal of Human-Computer Studies, 158*, Article 102744. https://doi.org/10.1016/j.ijhcs.2021.102744

Arnett, J. J., Offer, D., & Fine, M. A. (1997). Reckless driving in adolescence: 'State' and 'trait' factors. *Accident Analysis & Prevention, 29*, 57–63. https://doi.org/10.1016/S0001-4575(97)87007-8

Banks, V. A., Eriksson, A., O'Donoghue, J., & Stanton, N. A. (2018). Is partially automated driving a bad idea? Observations from an on-road study. *Applied Ergonomics, 68*, 138–145. https://doi.org/10.1016/j.apergo.2017.11.010

Beggiato, M., & Krems, J. F. (2013). The evolution of mental model, trust and acceptance of adaptive cruise control in relation to initial information. *Transportation Research Part F: Traffic Psychology and Behaviour, 18*, 47–57. https://doi.org/10.1016/j.trf.2012.12.006

Bhaskara, A., Skinner, M., & Loft, S. (2020). Agent transparency: A review of current theory and evidence. *IEEE Transactions on Human-Machine Systems, 50*, 215–224.

Billings, D., Schaefer, K., Chen, J., & Hancock, P. (2012). *Human-robot interaction: Developing trust in robots.*. https://doi.org/10.1145/2157689.2157709

Blanco, M., Atwood, J., Vasquez, H. M., Trimble, T. E., Fitchett, V. L., Radlbeck, J., . . . Cullinane, B. (2015). *Human factors evaluation of level 2 and level 3 automated driving concepts*.

Campbell, J. L., Brown, J. L., Graving, J. S., Richard, C. M., Lichty, M. G., Bacon, L. P., . . . Sanquist, T. (2018). *Human factors design guidance for level 2 and level 3 automated driving concepts*.

Cao, J., Lin, L., Zhang, J., Zhang, L., Wang, Y., & Wang, J. (2021). The development and validation of the perceived safety of intelligent connected vehicles scale. *Accident Analysis & Prevention, 154*, Article 106092. https://doi.org/10.1016/j.aap.2021.106092

Chen, J. Y., Procci, K., Boyce, M., Wright, J., Garcia, A., & Barnes, M. (2014). Situation awareness-based agent transparency. *US Army Research Laboratory*, 1–29.

Chiou, E. K., & Lee, J. D. (2023). Trusting automation: Designing for responsivity and resilience. *Human Factors, 65*, 137–165.

Dixon, L. (2020). Autonowashing: The Greenwashing of Vehicle Automation. *Transportation research interdisciplinary perspectives, 5*, Article 100113. https://doi.org/10.1016/j.trip.2020.100113

Ekman, F., Johansson, M., & Sochor, J. (2017). Creating appropriate trust in automated vehicle systems: A framework for HMI design. *IEEE Transactions on Human-Machine Systems, 48*, 95–101.

Endsley, M. R. (2016). *Designing for situation awareness: An approach to user-centered design.* CRC Press.

Endsley, M. R., Bolté, B., & Jones, D. G. (2003). *Designing for situation awareness: An approach to user-centered design.* CRC Press.

Eysenck, H. J. (1965). Fact and fiction in psychology. *(No Title)*.

Frijda, N. H., Manstead, A. S. R., & Bem, S. (2000). The influence of emotions on beliefs. In A. S. R. Manstead, N. H. Frijda, & S. Bem (Eds.), *Emotions and Beliefs: How Feelings Influence Thoughts* (pp. 1–9). Cambridge: Cambridge University Press.

Glaser, B. G., Strauss, A. L., & Strutzel, E. (1968). The discovery of grounded theory; strategies for qualitative research. *Nursing Research, 17*, 364.

Goodall, N. (2023). Normalizing crash risk of partially automated vehicles under sparse data. *Journal of Transportation Safety & Security, 1–17*. https://doi.org/10.1080/19439962.2023.2178566

Haghzare, S., Campos, J. L., Bak, K., & Mihailidis, A. (2021). Older adults' acceptance of fully automated vehicles: Effects of exposure, driving style, age, and driving conditions. *Accident Analysis & Prevention, 150*, Article 105919. https://doi.org/10.1016/j.aap.2020.105919

Haghzare, S., Stasiulis, E., Delfi, G., Mohamud, H., Rapoport, M. J., Naglie, G., & Campos, J. L. (2022). Automated vehicles for people with dementia: A "tremendous potential" that "has ways to go" – Reports of a Qualitative Study. *The Gerontologist, 63*, 140–154. https://doi.org/10.1093/geront/gnac115

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., & Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human Factors, 53*, 517–527.

Hardman, S., Chakraborty, D., & Kohn, E. (2021). A quantitative investigation into the impact of partially automated vehicles on vehicle miles travelled in California.

Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*, 904–908. https://doi.org/10.1177/154193120605000909

He, X., Stapel, J., Wang, M., & Happee, R. (2022). Modelling perceived risk and trust in driving automation reacting to merging and braking vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour, 86*, 178–195.

Hoff, K., & Bashir, M. (2013). A theoretical model for trust in automated systems. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 115–120).

Hoff, K. A., & Bashir, M. (2015). Trust in automation: Integrating empirical evidence on factors that influence trust. *Human Factors, 57*, 407–434.

Hoyle, R. H., Stephenson, M. T., Palmgreen, P., Lorch, E. P., & Donohew, R. L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality and Individual Differences, 32*, 401–414. https://doi.org/10.1016/S0191-8869(01)00032-0

Jabareen, Y. (2009). Building a conceptual framework: Philosophy, definitions, and procedure. *International Journal of Qualitative Methods, 8*, 49–62.

Johnson-George, C., & Swap, W. C. (1982). Measurement of specific interpersonal trust: Construction and validation of a scale to assess trust in a specific other. *Journal of Personality and Social Psychology, 43*, 1306.

Jonah, B. A. (1997). Sensation seeking and risky driving: A review and synthesis of the literature. *Accident Analysis & Prevention, 29*, 651–665. https://doi.org/10.1016/S0001-4575(97)00017-1

Kazi, T. A., Stanton, N. A., Walker, G. H., & Young, M. S. (2007). Designer driving: Drivers' conceptual models and level of trust in adaptive cruise control. *International Journal of Vehicle Design, 45*, 339–360.

Kim, H., Song, M., & Doerzaph, Z. (2021). *Is driving automation used as intended?* Transportation Research Record: Real-world use of partially automated driving systems and their safety consequences, 03611981211027150.

Kohn, S. C., de Visser, E. J., Wiese, E., Lee, Y.-C., & Shaw, T. H. (2021). Measurement of Trust in Automation: A Narrative Review and Reference Guide. *Frontiers in Psychology, 12*. https://doi.org/10.3389/fpsyg.2021.604977

Körber, M. (2018). Theoretical considerations and development of a questionnaire to measure trust in automation. *Paper presented at the Congress of the International Ergonomics Association*.

Lajunen, T. (2001). Personality and accident liability: Are extraversion, neuroticism and psychoticism related to traffic and occupational fatalities? *Personality and Individual Differences, 31*, 1365–1373. https://doi.org/10.1016/S0191-8869(00)00230-0

Lazarus, R. S. (1991a). Cognition and motivation in emotion. *American psychologist, 46*, 352.

Lazarus, R. S. (1991b). *Emotion and adaptation.* Oxford University Press.

Lee, J., Abe, G., Sato, K., & Itoh, M. (2021). Developing human-machine trust: Impacts of prior instruction and automation failure on driver trust in partially automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour, 81*, 384–395. https://doi.org/10.1016/j.trf.2021.06.013

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics, 35*, 1243–1270.

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*, 50–80.

Lenneman, J., Mangus, L., Jenness, J., Petraglia, E. (2020). Delineating Clusters of Learners for Driver Assistance Technologies. Paper presented at the HCI International 2020–Late Breaking Posters: 22nd International Conference, HCII. (2020). Copenhagen, Denmark, July 19–24, 2020. *Proceedings, Part, II*, 22.

Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. *Social Forces, 63*, 967–985.

Ljubi, K., & Groznik, A. (2023). Role played by social factors and privacy concerns in autonomous vehicle adoption. *Transport Policy, 132*, 1–15. https://doi.org/10.1016/j.tranpol.2022.12.013

Lynn, R., & Hampson, S. L. (1975). National differences in extraversion and neuroticism. *British Journal of Social and Clinical Psychology, 14*, 223–240.

Maarten Schraagen, J., Kerwien Lopez, S., Schneider, C., Schneider, V., Tönjes, S., & Wiechmann, E. (2021). The role of transparency and explainability in automated systems. *Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting*.

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., & Häussler, T. (2021). *Applying LDA topic modeling in communication research: Toward a valid and reliable methodology. In Computational methods for communication science* (pp. 13–38). Routledge.

Mason, J., Gaspar, J., Carney, C., Kim, J., Domeyer, J., Lenneman, J., & McGehee, D. (2023). Do Drivers Require Education After Over-The-Air Adas Updates?. *Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting*.

Matthews, G., & Desmond, P. A. (2000). *Stress and driving performance: Implications for design and training. In Stress, workload, and fatigue* (pp. 211–232). CRC Press.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of management review, 20*, 709–734.

McWilliams, T., & Ward, N. (2021). Underload on the road: Measuring vigilance decrements during partially automated driving. *Frontiers in psychology, 12*, Article 631364.

Mooradian, T., Renzl, B., & Matzler, K. (2006). Who trusts? Personality, trust and knowledge sharing. *Management learning, 37*, 523–540.

Mueller, A. S., Cicchino, J. B., & Calvanelli, J. V. (2023). Habits, attitudes, and expectations of regular users of partial driving automation systems. *Journal of Safety Research*. https://doi.org/10.1016/j.jsr.2023.10.015

Mueller, A. S., Reagan, I. J., & Cicchino, J. B. (2021). Addressing driver disengagement and proper system use: Human factors recommendations for level 2 driving automation design. *Journal of Cognitive Engineering and Decision Making, 15*, 3–27.

Muir, B. M. (1994). Trust in automation: Part I. Theoretical issues in the study of trust and human intervention in automated systems. *Ergonomics, 37*, 1905–1922.

Nordhoff, S. (2024). A conceptual framework for automation disengagements. *Scientific Reports, 14*, 8654. https://doi.org/10.1038/s41598-024-57882-6

Nordhoff, S., Lee, J. D., Calvert, S. C., Berge, S., Hagenzieker, M., & Happee, R. (2023). (Mis-)use of standard Autopilot and Full Self-Driving (FSD) Beta: Results from interviews with users of Tesla's FSD Beta. *Frontiers in Psychology, 14*. https://doi.org/10.3389/fpsyg.2023.1101520

Noy, I. Y., Shinar, D., & Horrey, W. J. (2018). Automated driving: Safety blind spots. *Safety Science, 102*, 68–78. https://doi.org/10.1016/j.ssci.2017.07.018

Özkan, T., & Lajunen, T. (2005). Multidimensional Traffic Locus of Control Scale (T-LOC): Factor structure and relationship to risky driving. *Personality and Individual Differences, 38*, 533–545. https://doi.org/10.1016/j.paid.2004.05.007

Özkan, T., & Lajunen, T. (2007). The role of personality, culture, and economy in unintentional fatalities: An aggregated level analysis. *Personality and Individual Differences, 43*, 519–530. https://doi.org/10.1016/j.paid.2006.12.020

Paardekooper, J.-P., Comi, M., Grappiolo, C., Snijders, R., van Vught, W., & Beekelaar, R. (2021). A Hybrid-AI Approach for Competence Assessment of Automated Driving functions. *Paper presented at the SafeAI@ AAAI*.

Palmeiro, A. R., van der Kint, S., Hagenzieker, M., Van Schagen, I., & De Winter, J. (2018). Cyclists' expectations when encountering automated vehicles: Results of an international photo-based questionnaire. *Paper presented at the ICSC 2018: 7th International Cycling Safety Conference*.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance Consequences of Automation-Induced 'Complacency'. *The International Journal of Aviation Psychology, 3*, 1–23. https://doi.org/10.1207/s15327108ijap0301_1

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors, 39*, 230–253.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. *Paper presented at the Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*.

Pennington, J., Socher, R., Manning, C. D. GloVe: Global Vectors for Word Representation. Retrieved from https://nlp.stanford.edu/projects/glove/.

Ploessl, F., Just, T., & Wehrheim, L. (2021). Cyclicity of real estate-related trends: Topic modelling and sentiment analysis on German real estate news. *Journal of European Real Estate Research, 14*, 381–400.

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49*, 95.

Rudin-Brown, C. M., & Noy, Y. I. (2002). Investigation of Behavioral Adaptation to Lane Departure Warnings. *Transportation Research Record, 1803*, 30–37. https://doi.org/10.3141/1803-05

Russell, S. M., Atwood, J., & McLaughlin, S. B. (2021). *Driver expectations for system control errors, driver engagement, and crash avoidance in level 2 driving automation systems*.

Schaefer, K. E., Billings, D. R., Szalma, J. L., Adams, J. K., Sanders, T. L., Chen, J. Y., & Hancock, P. A. (2014). *A meta-analysis of factors influencing the development of trust in automation: Implications for human-robot interaction*.

Schaefer, K. E., Chen, J. Y., Szalma, J. L., & Hancock, P. A. (2016). A meta-analysis of factors influencing the development of trust in automation: Implications for understanding autonomy in future systems. *Human Factors, 58*, 377–400.

Schaefer, K. E., Perelman, B. S., Gremillion, G. M., Marathe, A. R., & Metcalfe, J. S. (2021). Chapter 12 - A roadmap for developing team trust metrics for human-autonomy teams. In C. S. Nam, & J. B. Lyons (Eds.), *Trust in Human-Robot Interaction* (pp. 261–300). Academic Press.

Stocker, A. (2022). *Driver Trust in Automated Driving Systems.* Paper presented at the 30th European Conference on Information Systems: ECIS 2022.

Thushari, P. D., Aggarwal, N., Vajrobol, V., Saxena, G. J., Singh, S., & Pundir, A. (2023). Identifying discernible indications of psychological well-being using ML: Explainable AI in reddit social media interactions. *Social Network Analysis and Mining, 13*, 141. https://doi.org/10.1007/s13278-023-01145-1

Walker, F., Forster, Y., Hergeth, S., Kraus, J., Payre, W., Wintersberger, P., & Martens, M. (2023). Trust in automated vehicles: Constructs, psychological processes, and assessment. *Front Psychol, 14*, 1279271. https://doi.org/10.3389/fpsyg.2023.1279271

Watanabe, K., & Zhou, Y. (2022). Theory-driven analysis of large corpora: Semisupervised topic classification of the UN speeches. *Social Science Computer Review, 40*, 346–366.

Waung, M., McAuslan, P., & Lakshmanan, S. (2021). Trust and intention to use autonomous vehicles: Manufacturer focus and passenger control. *Transportation Research Part F: Traffic Psychology and Behaviour, 80*, 328–340.

Weller, G., & Schlag, B. (2004). Verhaltensadaptation nach Einführung von Fahrerassistenzsystemen. *Verkehrspsychologie. Mobilität - Sicherheit - Fahrerassistenz*, 351–370.

Wicki, M., Brückmann, G., Quoss, F., & Bernauer, T. (2023). What do we really know about the acceptance of battery electric vehicles?–Turns out, not much. *Transport reviews, 43*, 62–87.

Xu, Z., Zhang, K., Min, H., Wang, Z., Zhao, X., & Liu, P. (2018). What drives people to accept automated vehicles? Findings from a field experiment. *Transportation Research Part C: Emerging Technologies, 95*, 320–334.

Zhang, T., Tao, D., Qu, X., Zhang, X., Zeng, J., Zhu, H., & Zhu, H. (2020). Automated vehicle acceptance in China: Social influence and initial trust are key determinants. *Transportation Research Part C: Emerging Technologies, 112*, 220–233. https://doi.org/10.1016/j.trc.2020.01.027

Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking.* Cambridge University Press.