# Work with AI and Work for AI: Autonomous Vehicle Safety Drivers' Lived Experiences

Mengdi Chu
Institute for AI Industry Research,
Tsinghua University
Beijing, China
mengdichu@outlook.com

Keyu Zong
Institute for AI Industry Research,
Tsinghua University
Beijing, China
zongky@mails.ccnu.edu.cn

Xin Shu
Institute for AI Industry Research,
Tsinghua University
Beijing, China
mr_comfort@163.com

Jiangtao Gong*
Institute for AI Industry Research,
Tsinghua University
Beijing, China
gongjiangtao2@gmail.com

Zhicong Lu
Department of Computer Science,
City University of Hong Kong
Hong Kong, China
zhiconlu@cityu.edu.hk

Kaimin Guo
Institute for AI Industry Research,
Tsinghua University
Beijing, China
15536606362@163.com

Xinyi Dai
Institute for AI Industry Research,
Tsinghua University
Beijing, China
daixinyi53@foxmail.com

Guyue Zhou
Institute for AI Industry Research,
Tsinghua University
Beijing, China
zhouguyue@air.tsinghua.edu.cn

## ABSTRACT

The development of Autonomous Vehicle (AV) has created a novel job, the safety driver, recruited from experienced drivers to supervise and operate AV in numerous driving missions. Safety drivers usually work with non-perfect AV in high-risk real-world traffic environments for road testing tasks. However, this group of workers is under-explored in the HCI community. To fill this gap, we conducted semi-structured interviews with 26 safety drivers. Our results present how safety drivers cope with defective algorithms and shape and calibrate their perceptions while working with AV. We found that, as front-line workers, safety drivers are forced to take risks accumulated from the AV industry upstream and are also confronting restricted self-development in working for AV development. We contribute the first empirical evidence of the lived experience of safety drivers, the first passengers in the development of AV, and also the grassroots workers for AV, which can shed light on future human-AI interaction research.

## CCS CONCEPTS

• **Human-centered computing** → **Empirical studies in collaborative and social computing**.

---

*Corresponding author

---

## KEYWORDS

autonomous driving, human-AI interaction, AI perception, AI labor, responsible AI

## 1 INTRODUCTION

The arrival of autonomous vehicle technologies is expected to revolutionize human daily transportation and promises to enhance road safety, comfort, and mobility. This rapidly growing field has established a thriving industry in just a matter of a decade [8, 78]. "Safety Driver" is born of this trend. For technical, legislative, and ethical reasons, fully automated vehicles have not yet been widely implemented, human supervision of AV will still be needed for a long period [36, 54]. Hence, AV companies recruit experienced drivers as safety drivers to supervise and operate autonomous cars to ensure safety and conformity. Safety drivers typically have close and long-term interactions with autonomous vehicles in real-world scenarios. Understanding their practices, experiences, and challenges when working with highly automated systems can offer a glimpse into the upcoming autonomous society and inspire research on human-AI interaction. However, this group is under-explored in the HCI community. Most studies on human-autonomous vehicle interaction are based on laboratory environments or short-term observations that may be isolated from real-world practices and few studies have investigated the day-to-day interactions between drivers and highly automated systems. To fill this gap, we conducted semi-structured interviews with 26 safety drivers.

We explored the following research questions in this study:

(1) What are the safety drivers' work practices?
(2) How do safety drivers perceive, understand, and partner with AV technologies in working with AI?
(3) What are the experiences and challenges faced by safety drivers in working for AI?

We drew a picture of safety drivers' lived experiences and presented how safety drivers perceive, understand, and work with AV technologies:

- We examined how individuals with limited knowledge of AV form and adjust their perception of AV, and identified the factors that shape their perception.
- We investigated the transition of control between the safety driver and the autonomous system and uncovered tensions between organizational tendencies and individual tendencies.
- We found safety drivers' takeover decision-making characteristics in high-risk emergency situations.
- We revealed the learning preferences of safety drivers: practices than theories; try and fail than smooth processes; tangible and visible than abstract and invisible; and interaction with colleagues than taught lessons.

We also presented their work experiences, challenges, and well-beings while working for AV industries:

- We introduced their distinctive work experiences brought about by AV.
- We dove into the ambiguous responsibility allocation and moral dilemmas in their work.
- We revealed the well-being challenges faced by safety drivers: assuming risks generated by the upstream AV industries, limited opportunities for personal growth, and marginalization.

We compared our results with previous studies and explored the potential to enhance human-vehicle partnerships and improve worker experiences. Our research contributes the first empirical evidence of long-term interactions with autonomous vehicles in real-world scenarios. Furthermore, as a real world example of human-AI interactions, this study can serve as an analogy for broader human-AI research that involves real-world implications and provide insights for future human-AI studies.

## 2 BACKGROUND

The Society of Automotive Engineers (SAE) established defined 6 levels of AV to distinguish the responsibilities between the driver and the vehicle, ranging from Level 0 (no automation) to Level 5 (full automation) [14]. Level 1 and Level 2 are commonly referred to as "advanced driver-assistance systems" (ADAS), in which the human driver is in charge of driving tasks and receives assistance from the automation system; Level 3 is able to carry out critical driving tasks under certain conditions, and human drivers are supposed to continuously monitor and take control at all times; Level 4 automation is capable of performing all dynamic driving tasks under specific conditions, and a human driver may take control as necessary; At the highest level, Level 5 is able to perform all driving tasks under any circumstances without any intervention from the driver [24, 78]. The fully automated system without any

human intervention is the desired future of AV. However, this is not attainable in the short term, for technical, legislative, and ethical reasons [36, 54]. Human supervision for automated vehicles is still indispensable.

Generally, autonomous vehicles must undergo thorough public road testing to assess their viability and safety before being made available to the public [1]. Regions around the world have enacted laws and regulations to guide AV real-road testing. Although regulations vary among countries, most of them require that each autonomous vehicle be equipped with at least one human supervisor, who is responsible for monitoring the vehicle to ensure that the car is driving safely and adhering to traffic laws, and they must be ready to take over the autonomous vehicle if necessary [86].

Up to the time of this study conducted, under Chinese legal provisions, AV organizations have to ensure the car is under the supervision of human drivers in public road testing, regardless of the level of automation being tested. Driven by the fast-paced development of AV technology and the desire to trial it in real-world conditions, more and more AV organizations began conducting tests on public roads, which gave birth to a new occupation, "Safety Driver." AV organizations recruit experienced drivers to supervise and operate AV for numerous driving missions to ensure vehicle safety and conformity on public roads. Safety drivers usually get along with AV closely and long-termly in real-world environments. To some extent, they are the first passengers in the development of AV, as well as the future workers in the upcoming automated society.

## 3 RELATED WORK

### 3.1 Human-AI Partnership in Highly Automated Vehicles

Recent advances in AI technologies have led to AI systems becoming more closely embedded in human society, in which people and AI interact in complex ways and work together to collaboratively solve problems and perform specific tasks [17, 67]. To unlock the potential synergies between humans and machines, a variety of topics in the research and application of human-AI partnerships have been explored, such as shared mental models, goal-alignment, and decision-making between humans and AI systems [18, 71, 74]. In highly automated vehicles, the AI system carries out most of the driving tasks that were previously performed by human drivers, however, due to many hurdles to the wide adoption of fully autonomous driving, ranging from reliability to liability issues, human supervision is still essential and the driver needs to partner with the vehicle to finish driving tasks, which changes the relationship between human and AI systems more cooperatively [79, 92]. Some studies are being performed to explore human-AI collaboration in autonomous vehicle systems, focusing on trust calibration [21, 86, 96], situation awareness [43, 45, 94], take-over control [20, 32, 39] and Human-Machine Interaction (HMI) systems [13, 28, 90]. Shahrdar et al. [72] discussed the misuse of automation caused by distrust and over-trust; Takács et al. [78] summarized the challenges in supervising AV: limited human drivers' performance in terms of accuracy, time delay, and complexity, incorrect environment awareness, untimely situation assessment, lack of traffic information, and distraction by non-driving activities;

Baltzer et al. [7] presented the conflicts of control authority and responsibility distribution and the challenges of communication and state alignment in the cooperation between humans and automated vehicles. Researchers are actively exploring approaches to improving human-vehicle partnerships, including increasing the explainability [41] and transparency [56] of AV systems to enhance the calibration of drivers' trusts [86], perceptions [33], and mental models [88], and inventing novel HMI to improve drivers' attention and situation awareness [25, 46, 95], etc. Previous studies provided valuable human-vehicle interaction insights, while most of them were built on controlled laboratory environments or short-term observation, which may be isolated from long-term real-world practice. This study aims to fill this gap and investigates the in-the-wild human-AI partnerships in highly autonomous vehicles from the perspective of safety drivers.

## 3.2 From Driver Experiences to Autonomy Experiences

In the field of autonomous driving, understanding the experiences of human drivers - an important part of future traffic by working alongside automation systems - will contribute to future research in human-autonomy interactions and challenges posed by automation [2, 63]. Previous studies have explored drivers' experiences, perspectives, attitudes, and acceptance of autonomous systems. Karvonen et al. [37] investigated drivers' interactions with the automated metro system in Helsinki by conducting observations and interviews. Their findings identified the challenges faced by the drivers, including the demands for dynamic, complex, and uncertain control, the risk of decision-making in exceptional situations, and the monotonous work routine, and also shed light on the importance of considering the human factor in the design and operation of highly automated transportation systems. Yang et al. [93] presented truck drivers' on-the-road experience and subjective acceptance of using Cooperative Adaptive Cruise Control (CACC) based on their 160-mile driving experiment in Northern California and identified the factors influencing their acceptance and usage of CACC, such as road environments, traffic conditions, individual differences, etc. Lee et al. [42] conducted fieldwork involving six participants who rode in a prototype autonomous car on real roads for six days to investigate their experience with autonomous vehicles and identify factors that significantly influence passengers' trust in autonomous vehicle including lack of information, unpredictability and value misalignment, etc. Much of the current empirical research on automated driving experiences has focused on automated mass transit and trucks, and the drivers in these studies typically lack practical experience with day-to-day use of automated systems. There is a lack of empirical research on the actual experience of drivers of highly automated passenger cars, which is seen as one of the most likely autonomous vehicles to be widely used by the general public [66]. Moreover, the drivers of autonomous vehicles are also the users of AI systems. Their experiences with AI systems are important to consider in the development and deployment of AI technologies. Some studies investigated the end-users' subjective perceptions and folk minds of algorithms to gain insights into refining human-AI interactions and improving user understanding and trust of AI [47, 75]. Understanding non-technical users' experiences

and perspectives about AI could serve as a valuable heuristic cue to improve the comprehensibility, accessibility, and applicability of AI systems to a much broader group.

## 3.3 AI workers and Socio-Technical HCI

The deployment of emerging technologies, including those related to autonomous vehicles, has significantly impacted the nature of work and the power and social dynamics within workplaces. This shift in technology and work practices requires a rethinking of the relationship between emerging technologies and human workers [5, 10]. Manyika et al. advocated than compared to discuss whether jobs will be lost, it was important to evaluate how work will change due to the increased interactions of human and autonomous systems [52]. Baltrusch et al. [6] investigated the impacts of automation technologies on work quality and workers' well-being and identified four factors: cognitive workload, collaboration fluency, trust, and acceptance and satisfaction. Bhoopalam et al. [40] explored the truck drivers' perspectives of autonomous vehicle technologies through conducting focus groups. This study reported the concerns of one of the key stakeholder groups in the transition to AV technology and emphasized the need for careful consideration of the impact on workers and the development of strategies to support them during this transition. Automation technologies may change human workers from operators to more supervisory roles [91], promote deskilling for many workers and a need for new skills [23], create new occupations and opportunities [86], or increase the marginalization and precariousness of low-skilled workers [81]. When developing and deploying new technologies, researchers are supposed to adopt interdisciplinary approaches to form a socio-technical concept that considers not only technical issues, but also ethical and social implications, which might reveal important design characteristics for the integration of technology into human society [11, 60] Safety driver as the new occupation created by AV industries and seldom documented yet. Learning their work practices could help us foresee how AV technology will be embedded in our society.

## 4 METHODS

In order to gain a deep understanding of safety drivers' work practices and experiences, we conducted semi-structured interviews with 26 safety drivers in China from March to July 2022, utilizing further in-depth probing and detailed inquiry [69]. All the interviews were conducted remotely due to the COVID-19 pandemic. This study was approved by the author's organization's ethics committee.

*Participants Recruitment*. We recruited participants through professional communities, social platforms, and personal contacts, using snowball and purposeful sampling [65]. The sampling process was iterative until saturation was reached. To gain a more comprehensive picture of the experiences of safety drivers, we intentionally oversampled safety drivers from different companies and female safety drivers that we would not have otherwise. Each participant was given a compensation of 30 USD as a token of appreciation. We recruited 26 participants. Their ages ranged from 25 to 44 years old. 24 were male, and 2 were female. Their years as safety drivers ranged from 0.5 years to 5.5 years. They came from

**Table 1: Participants Information**

| Demographic information | Participant counts |
| --- | --- |
| Age | 20-25 years (3), 25-30 years (7), 30-35 years (9), 35-40 years (5), >= 40 years (2) |
| Gender | Male(24), Female(2) |
| Education level | Middle school(2), High school(12), Junior college(7), Bachelor's (5) |
| Years of being safety drivers | <1 year (4), 1-3 years (12), 3-5 years (8), >5 years (2) |
| AV technologies worked with | L3 (5), L4 (12), L3 and L4 (9) |
| Number of AV companies worked in | 1 company (15), 2 companies (9), 3 companies (1), >3 companies (1) |
| Employment Status | Employed (17), Dimission (9) |

8 AV companies in China, and 11 of them had work experience at multiple companies.

All participants had experience with highly autonomous vehicles of L3 or L4. According to the interviews, both L3 and L4 AV systems were able to perform most driving tasks on urban roads and operate independently. When encountering a situation challenging to handle, the L3 AV system would actively hand over control to the safety driver, whereas the L4 AV system would enter a safe state, as defined by the system (e.g., pull over), and would not transfer control to the human actively instead of waiting for the driver to take over passively. Since neither L3 nor L4 autonomous vehicles were able to perfectly handle all situations on public roads, it was the primary responsibility of the safety driver to assess the risk level of the current driving scenario in real-time, determine whether the autonomous vehicle could handle it alone, and actively take over control of the vehicle before the risk occurred.

Table 1 presents detailed demographic information for each interviewee. To preserve an additional layer of anonymity and prevent participants from being identified within their workplaces, we omitted information about their companies and blurred their exact age and years of working experience.

*Procedures.* Before the interview, two authors conducted a field visit and observed safety drivers' workflow. This served the purpose of familiarizing the authors with the safety drivers' work practices, collecting background information, and guiding the outline of the semi-structured interviews. The key sections of the interview included: (1) participant backgrounds; (2) understanding their work practices; (3) understanding their perceptions of AV; (4) understanding their partnerships with AV; and (5) understanding their working experiences and well-being. The average duration of each interview was approximately one hour. Each participant was interviewed individually by two researchers using online voice communication software, and all interviews were audio-recorded.

*Data Analysis.* Our data consisted of 29 hours of audio recordings . Firstly, three authors transcribed the interviews verbatim and examined the transcripts. We then adopted a thematic analysis approach[50] to analyze the 26 transcripts. Each participant's data was qualitatively coded by two or three researchers[53]. The authors collaboratively analyzed the codes and grouped them into themes, and refined the relevant themes in relation to the research questions. At the end, we categorized our resultant codes and themes into three main sections to reflect our findings: (1) work

practices of safety drivers; (2) empirical information about the perceptions and partnerships of safe drivers with AV; and (3) safety drivers' work experiences and well-being in working for AV.

*Research ethics.* Before commencing the work, this study was approved by the ethics committee of the authors' organization. We obtained informed consent from each participant, and participants had the option to decline to answer any questions and terminate the interview at any time. To preserve participant confidentiality, all personally identifiable information was removed from research files, and any identifying details were omitted when quoting participants, due to the sensitive nature of the research.

## 5 RESULTS

This section presents the results of our interviews with safety drivers. Section 5.1 describes the current work practices of safety drivers. Section 5.2 focuses on the human-vehicle partnerships and presents how safety drivers perceive, understand, and work with autonomous vehicles. Section 5.3 describes safety drivers' experiences, well-being, and challenges in working for autonomous vehicle industries.

### 5.1 Being a safety-driver

*Recruitments.* Self-driving companies generally hire experienced drivers to ensure that autonomous vehicles can complete driving tasks safely and are corrected in time to avoid risks and accidents. According to our qualitative data, all participants' companies valued extensive driving experience, excellent responsiveness, and good driving habits, which were considered the prerequisites for becoming a safety driver and the basic recruitment requirements. Most of the safety drivers had prior experience in one or more driving-related occupations, such as taxi drivers (N=8), driving instructors (N=2), chauffeurs (N=2), truck drivers (N=3), and freelance drivers on gig platforms (N=16). We also performed statistics on the total number of years of the professional driving experience (all professional experience related to driving, including working as full-time drivers, freelance drivers, and safety drivers) that each participant had up to the time of the interview, and the 26 participants had an average of 6.8 years of professional driving experience. Additionally, the majority of the safety drivers interviewed (N=24) were under the age of 40. Our participants responded that AV companies tended to hire younger drivers than older drivers if both meet the recruitment criteria. Although, we did not find solid facts and statistics to prove that younger safety drivers are better equipped to supervise autonomous vehicles compared to older drivers in

this study, the bias and stereotypes surrounding older workers in the labor market may result in AV companies preferring to hire younger safety drivers.

> "Safety drivers need to understand some basic AV algorithmic knowledge and technical principles, but older people are often seen as being less open to new things and slower to learn. I guess that's why our company would rather recruit and train younger drivers than older drivers." (P4)

> "You must have very sharp reflexes to monitor self-driving cars. Older safety drivers may not react as quickly as younger ones." (P17)

> "Working as a safety driver is also physically demanding. You need to be stuck in the car all day and maintain a high level of concentration for a long time. That takes a lot of energy and physical strength, and the younger may be more preferable to the older." (P2)

**Motivations**. Our research showed that safety drivers are usually low-income workers who depend on their driving skills but have little knowledge about autonomous vehicles. Similar to other blue-collar groups in developing countries, they often have limited access to employment and education [55, 76]. Only seven of our participants had a bachelor's degree. They relied on their driving skills and had few other options for making a living. Many participants expressed that being a safety driver was a worthwhile opportunity, and a sound choice for them, as it could provide a reasonable and stable salary and improve their standard of living to some extent. As P8 said:

> "I'm not sure what else I can do besides driving. Although working as a safety driver doesn't pay very well, it's enough to cover my needs. The company offers me comprehensive insurance and a housing fund, which is much better for me than driving on DIDI (a gig driver platform in China)."

Aside from the financial benefits, one of the main reasons they wanted to be safety drivers was their fascinations with and enthusiasms for AV. According to our interviews, most of the participants (N=24) did not have technological backgrounds. They had little AV knowledge but a lot of enthusiasms for AV industries. Consistent with studies of low-skilled workers in AI industries, they were more likely to be drawn to these industries by the promising future of emerging technologies [82, 85].

> "I love trying new things, so the chance to work with autonomous driving really gets me pumped!" (P2)

> "During the interview, the recruiter introduced the development mileage of autonomous driving to me. I thought it was amazing and wanted to get involved to witness its transformation. It would be a great opportunity to improve myself and learn some new skills." (P12)

**Training and assessment**. Meeting the recruitment criteria did not guarantee these workers would become safety drivers eventually. Before their employment confirmation get approved, they must undergo rigorous training and pass several rounds of assessments. According to our participants, the training mainly focused

on how to operate an autonomous vehicle safely and driving behavior norms, through theoretical and practical training and lasted a few weeks to months depending on their companies. The theoretical training typically covered fundamental AV knowledge, driving behavior codes, AV control methods, AV supervision precautions, accident treatment, etc. Most participants reported that their companies did not provide them with in-depth information regarding the workings of self-driving technologies. P15 mentioned:

> "In the company's perspective, it is enough for us to drive the car safely, and we do not need to know the technologies behind it very well. Also, since we aren't particularly skilled at learning technology, the company doesn't feel it's worth the effort to teach us."

For practical training, novice safety drivers typically operated AV under the supervision of coaches or senior safety drivers. The training scenarios progressed from simple proving grounds to complex public urban roads. Over several weeks, they developed their supervisory skills and became familiar with AV operation. There were multiple rounds of evaluation, ranging from theoretical exams to practical operations, with a focus on responsiveness and driving behavior. Those who didn't pass the test must be retrained or leave the job. *"We had to pass three exams in total, and these exams were very strict. About half of the people in our group failed,"* P15 said. Workers who successfully pass those assessments can eventually become safety drivers, starting their journeys to working with autonomous vehicles.

**Responsibilities**. Safety drivers were responsible for supervising and operating AV to finish road testing tasks and ensuring the vehicles' safety and conformity. According to our interviews, takeover, which typically includes braking, throttling up, and turning the steering wheel, was the most crucial operation for safety drivers to interfere with AV, and important not only for vehicles' safety but also for the research and development team to analyze system defects according to its records. Safety drivers should ensure both the safety of the vehicle and the reliability of the data produced by takeover. Hence, their takeover decision-making needs to strike a balance between safety and data quality. To achieve this, they need to form precise mental models to predict the behavior of autonomous vehicles and acquire accurate situational awareness of the driving environment. There were two types of road testing forms: "1 safety driver + 1 AV" and "1 safety driver + 1 AV + engineer(s)." Additionally, based on the different company's organizational structure, division of responsibilities, and tasks assigned, safety drivers might be required to perform other duties such as data recording, debugging assistance, and hardware maintenance.

**Performance Appraisal**. The performance appraisal systems and responsibility allocation regulations for safety drivers differed depending on the company's philosophy and policies. Common criteria for evaluating their performance included mileage driven, working hours, and accident rate. The accident rate was the most important criterion and was taken very seriously by AV companies. 8 participants reported that they would face punishments or even be fired by their companies for accidents.
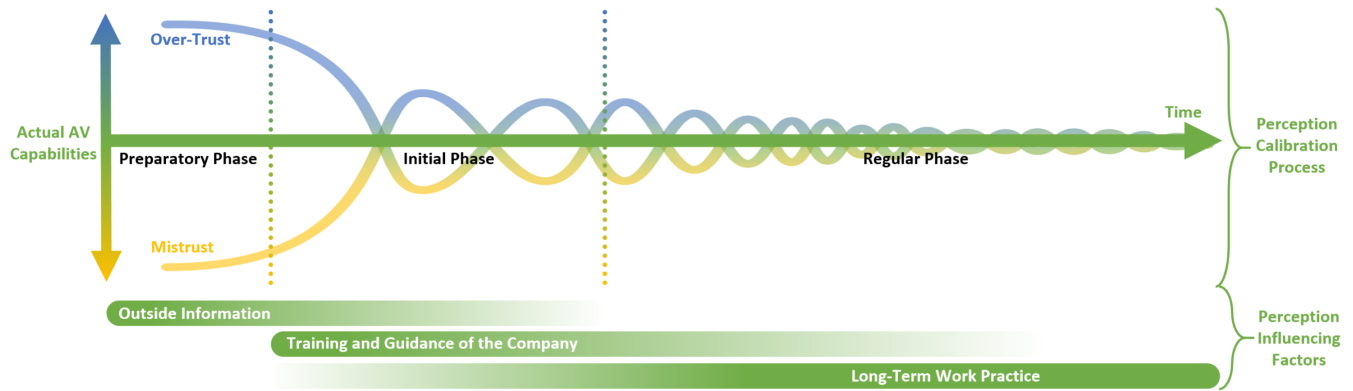
**Figure 1: The calibration process of safety drivers' perceptions of autonomous vehicles' capabilities.**

## 5.2 Working with AI: How safe drivers perceive, understand, and partner with AV

*5.2.1 Forming and calibrating perceptions of AV.* Based on the qualitative data, we identified factors that influence driver perceptions of AV capabilities and classified the calibration process into three stages: preparatory, initial, and regular, as shown in Figure 1. We found that during the preparatory phase, due to a lack of knowledge about AV, their perception was formed based on outside sources such as news media, journalists, social media, etc., which often contain incorrect information. And there was a higher likelihood of technophobia [9] and technopraise [35] among this group, which may result in over-trust or mistrust about AV capabilities. In the initial phase, as they gain access to AV and receive training and guidance from their companies, their perceptions will be quickly calibrated.

> *"At first, I was skeptical. I couldn't wrap my head around how a few tons of metal could drive on its own. I was even scared to get into the car at the start of my work. But after a few days, I started to feel more comfortable, and it exceeded my expectations."* (P10)

> *"I used to think self-driving cars were all they were cracked up to be, but after driving it during training, I found that it wasn't like what I had seen on TV. I found that it wasn't like what I had seen on TV. I just couldn't relax and trust it to drive itself."* (P26)

In the regular phase, safety drivers continuously calibrated their perceptions through their work practices, causing their mental models to converge with the actual capabilities of AV over time. *"It took me about a year to really get to know the car, and I feel like I'm getting better at making predictions now"*, P1 said.

We found that company-level factors such as training and guidance had a significant impact on safety drivers' perceptions in the initial phase. However, in the regular phase, their AV perceptions became more based on their hands-on practices than company-level factors. Many participants reported that the theoretical pipe-lined training they received didn't help much with calibrating their mental models, and it was difficult to apply that theoretical knowledge into real-world driving practices, especially in high-risk and emergency situations. Instead, they found that constantly trialing and erroring in their work practices was a better way for them to explore the boundaries of AV capabilities and calibrate their perceptions accurately. Also, they reported that the lessons gained in these processes would be more impressive.

> *"Although the company has informed us that the radar often misses low objects, if you suddenly encounter obstacles in the road, it's tough to act on that information right then and there. But after being startled, you can remember this lesson very well and handle it better next time."* (P12)

Additionally, participants reported that after the "novice period," their companies usually paid less attention to the accuracy of their mental models about AV and didn't provide adequate training and assessments in a timely manner. There was also a lack of official methods for continuously calibrating their perceptions while working with AV on a long-term basis. With each update to the algorithms, system iteration, or hardware change, safety drivers had to form new mental models. However, they didn't know what the "standard answer" was. They had to test and adjust their perceptions in high-risk environments, which resulted in their being potentially exposed to safety risks resulting from misperceptions, and they had to assume responsibilities for those risks.

> *"The engineers usually only give us a rough idea of version characteristics when the version is updated, and they don't know when the car might have a malfunction. We must concentrate our attention while driving to try and identify any potential problems with the car."* (P22)

*5.2.2 Takeover in real world.* According to the interviews, takeover was the most important operation for safety drivers to interfere with AV, including braking, throttling up, and turning the steering wheel, which often took place in risky and unexpected situations and would be recorded by the AV system for problem analysis. For the majority of the systems our participants worked with (N=23), once the driver took control of the vehicle, the automation system would shut down and transfer control to the human.
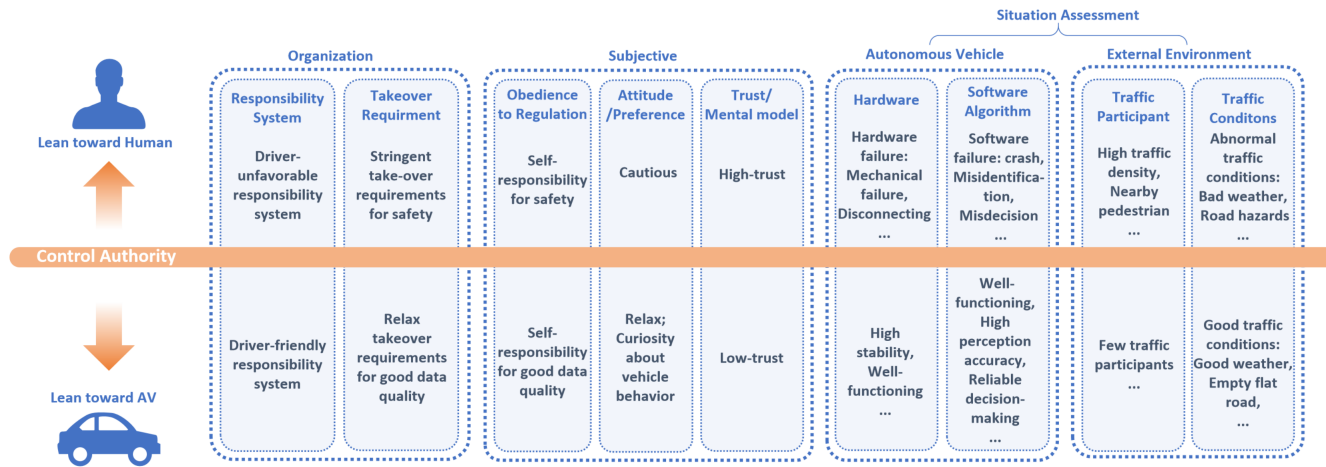
**Figure 2: Influencing factors of control authority between the human driver and the autonomous vehicle.**

From the qualitative data, we identified three sets of factors that impacted safety drivers' takeover decision-making and vehicle control authority: organizational factors, personal tendencies, and real-time situations (including the autonomous vehicle and the external environment), as shown in Figure 2.

We found that in low-risk and non-emergency situations, the safety drivers were more likely to be influenced by company factors when making takeover decisions. But in emergency situations, participants expressed a desire to take control of the vehicle and made more subjective and intuitive takeover decisions, which may go against the company's requirements, guidelines, and advocacy.

***Organizational tendencies vs. individual tendencies.*** Many participants reported that their companies usually guide the driver's takeover criteria based on corporate strategies. Companies that prioritized efficiency in testing and the quality of takeover data were more likely to ask their safety drivers to relax their takeover criteria. P23 mentioned,

> *"My company wants us to be less sensitive to takeovers so we can better know the reaction of AV. If you feel something wrong and take over it immediately, you won't know what its abnormal reaction exactly is."*

While companies that prioritized test safety were more likely to ask their safety drivers to take over in advance. P19 said,

> *"My leader always emphasizes, 'if you feel something wrong, you must take-over promptly, and you should not tempt the reaction boundaries of the car, otherwise, you are risking the company's property and your personal safety,' so we rarely tempt the limits of the car."*

Moreover, safety drivers' takeover attitudes also fluctuated according to their subjective tendencies. P24 said,

> *"Occasionally, the car will act strangely, and if there are no other cars on the road, I will not take over control purposefully to see how the car will behave. Obviously, I shouldn't do that, and my actions may be recorded by the system. I only do this on a few rare occasions out of curiosity."*

Sometimes, there were conflicts between the company's tendencies and individual tendencies. In non-emergency situations, these tensions were not obvious. Most safety drivers tended to follow their companies' guidelines to adjust their takeover attitudes and tighten or relax their takeover criteria. As P8 mentioned, *"This is my responsibility to meet my company's requirements. I need to get my job done well."*

However, in high-risk emergencies, participants reported that they would disregard the company's guidelines and rely on their intuition due to concerns about the consequences of risk, such as their personal safety, liability for accidents, and traffic laws. They explained that they wanted to take control of the vehicle and felt that having more control would give them a greater sense of safety.

> *"If I feel danger, I will take over control. In that moment, I am unable to give any thought to the company's requirements."* (P12)

> *"After taking over and getting the wheel in my hands, I feel relieved and safer."* (P10)

Some participants said they had more faith in their own decisions and believed that human decisions were better than those made by machines, especially in emergency situations. It is worth mentioning that the majority of the participants (N=21) stated that they preferred to take control of the vehicle themselves in risky situations, even if the self-driving decision might be the correct one. As P16 stated,

> *"Even if it happens many times that the autopilot is right and my judgment is wrong, the next time there is a conflict between our decisions, I would still choose to make it listen to me. Although its ability is excellent, I don't think I can rest assured to give all of myself to it."*

***For real-time situation assessment.*** Most safety drivers (N=24) reported they needed to pay much more attention when supervising the self-driving vehicles compared to when they were driving manually. Extensive lab-experiment-based literature describes that high-level autonomous driving may bring distraction [16], loss of attention [59], decreased situation awareness [30], etc. to drivers.

However, few participants mentioned that their level of attention decreased when supervising AV, even for long periods of time. There were even some reported that their level of attention increased with driving time.

> *"There are so many unpredictable factors involved. When driving for a long time, I feel that my focus improves."* (P22)

> *"The more time you spend driving, the more tense you become."* (P1)

Safety drivers were required to take regular breaks, and their driving time typically did not exceed one hour at a time. This was perhaps one reason why there were no reports of decreased attention caused by long-time automation. Also, participants interviewed in this study were urban road safety drivers, who encountered driving scenarios usually with low monotony and high uncertainty. They had to deal with the uncertainty of self-driving hardware, software, and external environments and making quick decisions when something wrong. The unpredictability factors in the real world made them continuous high attentions. *"I need to pay more attention to the surroundings and detect them more frequently when monitoring AV. I will be very alert and my attention will be more focused,"* P7 said. Meanwhile, long periods of concentration also lead to fatigue in safety drivers' work. P2 said, *"It's very brain-intensive, and I feel very tired at the end of the day."*

Moreover, participants reported that, due to the fact that their personal safety and work performance were closely tied to the safety of autonomous cars in the real world, they were unable to let their guard down, which was different from the safety-guaranteed laboratory-based studies. Additionally, upon the company's request, the perception and decision-making of safety drivers needed to be paralleled with that of autonomous vehicles. Hence, reducing attention and handing over part of the driving perception tasks to AV would have been negligent for safety drivers. We inferred that one of the possible reasons for increased attention was that our participants were professional safety drivers, had passed related training, and would increase their attention proactively at work. But we did not exclude the possibility that the sensitivity of their work might have deterred them from speaking more critically and admitting their dereliction of duty.

*5.2.3 Cognitive Preferences and Characteristics of AV Technologies.* **Know little but want to know more.** The close interaction between safety drivers and AV made exposure to AV technologies in their daily work inevitable, such as learning codes to assist engineers in testing, detecting, and fixing minor AV problems, which reduced their sense of separation and awe about the technology while also stimulating their curiosity. Despite the fact that most our participants had limited knowledge of AV technologies and no technical background, they expressed a desire to learn about how AV works and AV technical principles. P14 said,

> *"I spend every day with this car, but I don't know what it's thinking. I want to know why it makes certain decisions and behaves the way it does, and I want to know the logic behind it."*

**Learning from interacting with engineers.** We found that most AV companies believed that it was sufficient for safety drivers just to be drivers for AV and did not invest in technology-related training for them. Participants reported that most of their understanding of AV technologies gained through their interactions with engineers. P13 mentioned,

> *"When testing with the engineer, he is too busy to handle all the tasks himself. He usually teaches me some codes so that I can assist him in tuning the program, and show me how to query the database and fix some bugs. I often ask him why certain things can be tuned in certain ways during the test, and he will explain what the code stands for and what else it can do."*

P17 mentioned that he would establish a good relationship with engineers who were willing to share their knowledge, in order to facilitate future consultation and learning.

> *"If I build a good relationship with the engineer, he may request that I be paired with the captain so that I can spend more time communicating with him and learn something."*

Although communication with engineers can help safety drivers understand technology to a certain extent, this knowledge transfer was limited, and the information obtained might be incomplete. Engineers were not always concerned with how much safety drivers had mastered or whether their knowledge mastered was accurate. *"The company will not teach us these, and the engineers only speak these to us briefly, so we can only understand the superficial aspects,"* P15 said.

**Invisible algorithms and visible hardware.** We found that most safety drivers tend to associate invisible algorithms with visible hardware. When we asked them about their understanding of AV, most answered by talking about their knowledge of AV hardware. P19 said,

> *"I don't know much about software algorithms, but I do understand some basic hardware concepts like lidar, millimeter-wave radar, cameras, and so on. I understand their perception range, parameters, and simple hardware debugging methods."*

We also found they would naturally associate the level of AV capabilities with the vehicle's hardware systems. When there was a problem with the AV, they would first assume it was a hardware issue, and then consider the possibility of an algorithm problem. *"If the car suddenly stalls, I will first check its radar, camera, and other sensors,"* P15 said. Participants also reported their greater interests in understanding hardware knowledge than software algorithms. They believed hardware knowledge was easier to comprehend and more useful in their daily work. P14 and P16 mentioned that they would like more training on hardware knowledge from the company. P10, P18, and P21 expressed their desire to switch to a career as hardware engineers.

> *"I prefer purely mechanical things."* (P14)

> *"I have difficulty understanding those red and green codes, but I'm more familiar with how the hardware works."* (P16)

## 5.3 Working for AI: Safety Drivers' Experiences and Well-being

*5.3.1 Work Experiences. From technology experiences to work experiences.*

Safety drivers need to work in close contact with AV in their daily tasks. In such a work context, they were the supervisor and also the experienced user of AV, and their happiness and sadness at workplaces were closely linked to the AV they worked for.

As supervisors, participants reported that they were able to gain a great sense of effectiveness when the issues they reported were resolved, and they could feel firsthand the progress of AV. *"When we report a problem and it is suddenly resolved after a few days, I feel a sense of satisfaction knowing that I was able to contribute to the improvement of the vehicle,"* P24 said. Some indicated that they feel neglected and unappreciated when they do not receive feedback or the issues they report are not resolved. *"We've reported this issue several times, but it still happens as soon as the car approaches this intersection. It's been a long time, and the problem has not been resolved. I feel like they just don't want to pay much attention to me,"* P15 said.

As users of AV, participants reported that they have positive experiences when the vehicles behave in accordance with human thoughts in their daily driving.

> *"I want it to go faster, and it does; I want it to slow down, and it does. It's as if it knows what you're thinking, and it feels like I've become integrated with it."* (P11)

Some indicated that the inconsistency between the vehicle's behavior and the human's intentions and expectations might trigger their negative experiences.

> *"It has its own thoughts, and I have mine. Sometimes it doesn't inform me, it just goes right by and it just lets me agitated."* (P21)

Some expressed disappointment that the limitations of algorithms sometimes override human intentions. *"I like the aggressive driving style, but the car just drives very conservatively,"* P23 said. Participants also mentioned they would prefer to work with the "perfect machine".

> *"I've worked for two different companies. The technology of the first company was not very developed, as the autonomous vehicle would frequently brake abruptly and cause conflicts with other road participants. When I finish my daily work, I'm often in a bad mood. But now, the technology of my current company is much more developed, and the autonomous vehicle is much smoother and more comfortable to ride. I feel much better now."* (P9)

***Technology mediates social interactions in the workplace.*** Autonomous driving technology also had an impact on the social interactions among safety drivers at workplaces. We found that AV technologies improved communication and collegiality among them. P9 mentioned that when she first started at the company, she was unfamiliar with everything. AV became an ice-breaker topic between her and her colleagues, hastening their acquaintance. Safety drivers were willing to discuss their guesses about AV with others and share problem solutions, which gave them a sense of

accomplishment and self-efficacy and strengthened their colleague relationships. *"We often discuss common problems together, such as navigation and perception issues, and we are willing to help each other,"* P11 said.

We also found that safety drivers' interactions with engineers significantly enhanced their desire to learn more about AV technologies and their self-efficacy, which in turn impacted their career planning.

> *"The engineers have taught me a lot, and I feel like I can do the same things they do. I'm thinking about becoming an engineer in the future."* (P21)

While we observed positive effects of technology on workplace interactions, we also found that technology hindered some safety drivers' self-expression to some extent. Some participants mentioned that their lack of technological knowledge made them hesitant to express their opinions. They tended to suppress their expression in certain situations because they feared being denied or ridiculed.

> *"When I encounter some simple problems that I don't know how to solve, I'm afraid to ask my leader because I don't want them to think I'm incompetent. I often ask my colleagues, but they don't always know the answers to my questions."* (P11)

> *"Sometimes I have some guesses about AV problems. I usually wait until I'm absolutely certain before sharing them with others. After all, I'm not a professional engineer, so I'm concerned that my ideas may be incorrect."* (P23)

***Real world driving experiences.*** When safety drivers drove the "strange-looking black box" on public roads, they attracted more attention from the outside world and had novel experiences with AV, which gave them the pride of being noticed, unnecessary annoyances, and risks caused by other traffic participants' curiosity and low acceptance of new things. Some participants expressed a very positive feeling about the curiosity and attention of the outside world caused by autonomous driving and felt proud and happy.

> *"When I'm waiting for the traffic light at the intersection, people often take pictures of the autonomous vehicle, which gives me the illusion that they are paying attention to me."* (P2)

Some also mentioned that the outside attention would cause them unnecessary distress.

> *"Because this vehicle is too conspicuous, the traffic police sometimes specifically spot-check me."* (P6)

> *"When the vehicle breaks down on the road, some passers-by may go near and take a look, and even stop to take a picture, which made me very embarrassed."* (P16)

Furthermore, participants stated that they had to accept the risk of low acceptance of AV by other road participants, such as malicious behavior caused by curiosity.

> *"There are many cars on the road that often deliberately come to a halt in front of you because they are curious how the self-driving car will react."* (P20)

*"A few years ago, the news media widely reported that AV would replace human drivers. During that time, I was often bullied by taxis while conducting road tests. But it's much better now, they have become accustomed to our vehicles."* (P1)

Moreover, due to their limited capabilities, autonomous vehicles may not be able to handle all road test conditions perfectly, which may result in exclusion by other road participants. P23 mentioned that his vehicle sometimes has to bear the consequences of mistakes made by other autonomous vehicles while driving:

*"The cars in our team often make mistakes on the road and affect others (road participants). Those drivers probably hold a grudge, and the next time they encounter our team's autonomous vehicles, they will deliberately bully them, even if the car they retaliate against is not the same one that disrupted them last time. I am often wrongly accused."*

### 5.3.2 Taking Risks Accumulated from upstream of AV Industry.
**Forced to expose themselves to accumulated risks from AV upstream.** During road tests, safety drivers, as the downstream workforce of the AV industry, were forced to face the accumulated risks from multiple upstream links, including algorithm development, hardware manufacturing, assembly, etc. Unlike traditional testers [29], they needed to verify and test AV in high-risk real-world environments, where a minor omission from other links could expose them to great risks. Based on the interviews, developers in other upstream links of AV were not required to participate in road tests like safety drivers. In such a workflow, these stakeholders were unable to predict precisely how the AV they developed would perform in the real world, which might reduce their sense of responsibility and increase the likelihood of negligence.

*"I remember a remote debugging session. The engineer told me the car had been fixed and let me test it. But when I tested it, I found that the problem had not been set up right, and I almost collided with an obstacle."* (P7)

*"There was a system version update, but the new version had many strange problems, and the engineers didn't know what was going on. We felt very unsafe while driving and had to pay very close attention."* (P14)

Although it was the duty of the safety drivers to look for defects and intervene in AV to prevent risks, relying solely on the safety drivers to identify accumulated problems and avoid accumulated risks is difficult and immoral.

P3 said, *"There are some unexpected situations where it's too late for you to takeover, and no one can change the outcome in such a short amount of time."* P7 expressed his concerns about safety, *"Every morning when I go out for work, I pray to come home safely,"* and thoughts about changing jobs.

### 5.3.3 Ambiguous Responsibility Assignments and Moral wrinkles.
**Responsibility assignments.** In addition to bearing the accumulated risks, safety drivers also faced ambiguous and undue assignments of responsibility. China's traffic laws did not consider self-driving cars as responsible entities. When safety drivers conducted road tests with AV, they were held legally responsible. Even though the majority of the time the car was being driven by AI and not the safety driver, the safety driver was still held accountable for the vehicle. For traffic violations, even those caused entirely by the self-driving system, the safety driver was still punished as the primary responsible party and their driver's license points would be discounted. A poor driving record would be recorded in the driver's file and might impact their driving qualifications or limit their opportunities for other driving-related jobs. P8 said,

*"If we receive a citation for breaking traffic laws because of the self-driving car, the company will compensate us with some money. However, there is no way to regain the license points that have been deducted, and we have to accept the loss."*

Participants said they do not have to take responsibility for passive accidents, but if the accident was caused by the fault of the safety driver or the autonomous vehicle, they may face consequences such as a warning or penalty from the company. Two-thirds of the safety drivers reported that the company's responsibility allocation system was not fair to them.

*"The company has a zero-tolerance policy for accidents. Because safety drivers are hired for the purpose of ensuring safety. If an accident occurs, it is considered the responsibility of the safety driver."* (P15)

Some participants reported that their companies implemented a shared responsibility system as a way to warn safety drivers not to have an accident.

*"The training of new safety drivers is conducted by old safety drivers, and if the new safety driver causes an accident, the old safety driver will take responsibility for the new safety driver."* (P21)

It is undeniable that this stringent responsibility system can improve the safety drivers' sense of responsibility at work, but it also infringes on their rights and leads to a negative work experience to some extent. There were also some companies that differentiated the responsibilities between human and autonomous vehicles.

*"If accidents happen while on autopilot mode, your responsibility may be reduced, but if you were distracted, fell asleep, drank water, used your phone, etc., that would be your responsibility."* (P25)

However, due to the lack of clarity on the boundaries of responsibility, safety drivers may take on more responsibility than they are supposed to.

*"In the moment before an accident, you must have taken control instinctively. It can be difficult to determine whether you were involved in causing the accident or not."* (P17)

*"Although it is said that the company will analyze whether the accident was caused by AV or the safety driver through monitoring data and system records, but then again, if there is an accident and you take over, then you may be responsible for involving the accident. If you don't take over, that means you have not performed the duties of the safety driver, and then you are still responsible."* (P13)

**Operator or passive observer.** As a new type of occupation, the current corporate systems, laws, and regulations are not fully

equipped to guide safety drivers in all situations, leading to ambiguities in road testing. As described in the previous section, some companies adopt a safety-driver-friendly allocation of collision responsibility. If a collision occurs while the autonomous system is running without any human interference, the company will assume full responsibility for the autonomous system they developed. However, this regulation raises another question of whether the safety driver should intervene in the AV system at the moment of an upcoming accident. Participants reported that they would not stand by and watch an accident happen to reduce their responsibility. They would instead instinctively take over the autonomous vehicle, even if their actions would not save anything.

> *"Takeover better than an accident."* (P19)

> *"I'll take over even if I can't save it. I can't watch the car go into an accident."* (P11)

> *"I wouldn't allow an accident to happen for fear of liability, but I don't know what the rest of my colleagues think."* (P8)

> *"Even though not intervening could absolve me of responsibility, there's no guarantee that the company wouldn't hold me accountable if an accident occurs. So I believe it's more important to prevent an accident from happening."* (P16)

**Collision leaded by avoid-collision.** According to our interviews, AV companies usually adopt very cautious algorithmic strategies to avoid active collisions. However, these overly conservative driving strategies made self-driving cars slow and easy to stop, increasing the chances of being rear-ended by other vehicles. Participants reported that they have to put in extra effort and take on additional risks because of the shortcomings of AV strategies.

> *"When the self-driving car falls short, it's up to us safety drivers to fill in the gaps. But sometimes the car's capabilities are just too limited and we are also very helpless. But after all, this is our job."* (P22)

> *"The safety driver must pay more attention to the rear of the vehicle to avoid being rear-ended, which requires a high level of attentiveness from the safety driver. Sometimes some novice safety drivers are not able to effectively consider the surroundings, so it is quite prone to accidents."* (P25)

*5.3.4 "I can see the future of AV, but not mine".* As described in Section 5.1, safety drivers typically relied on their driving skills for their livelihood and had limited career choices. Although they had the opportunity to be exposed to emerging technologies and witness changes in the autonomous driving industry as safety drivers, it was challenging for them to achieve career growth through this job.

> *"You can't learn anything just by staring at this car every day. I feel as if I've been extinguished after working for a long period of time as a safety driver."* (P9)

Participants expressed their concerns about their future developments.

> *"As long as you can drive, you can be a safety driver. There is nothing irreplaceable. Who knows, I might be laid off one day."* (P5)

They also reported an age crisis among safety drivers. Some companies set age requirements due to the high levels of endurance, sensitivity and responsiveness required for the job. Despite many studies [19, 34, 38] showed that there was no significant difference in takeover response ability between young and middle-aged people, the inherent impression can also expose this group to age discrimination. *"Younger drivers may have better response abilities, and the older drivers may be gradually eliminated from the company,"* P10 said.

Moreover, participants also recognized that AV technology is moving towards unmanned operation, and they feared that the role of safety driver may become an overstaffed position that is vulnerable to being eliminated. P18 said,

> *"Safety driver is only a transitional position in AV development. With the progress of AV technology, safety drivers may become obsolete. When that happens, I don't know what the company's plan is, and I am not sure about my future."*

P9 expressed his ambivalence between his expectations for the development of AV and his future development as a safety driver:

> *"I wish for the technology to progress faster to make my job easier, but I don't want it to develop very well because it may jeopardize my employment."*

## 6 DISCUSSION

### 6.1 From Human-Vehicle Partnerships to broader Human-AI Partnerships

*More effective training and experience transfer for non-technical lay users.* Designing effective training programs and experience transfer strategies can help ensure that the technology is used in a responsible and ethical manner. Proper training of the driver in taking over the vehicle and prompt decision making is imperative for ensuring safety. According to our interviews, although the process of understanding and learning AV technologies may vary for safety drivers based on their individual preferences, experiences, and circumstances, some common characteristics of experience acquisition have been observed in this study. These include: *practices over theories, try and fail over smooth processes, specific and visible over abstract and invisible, and interactions with colleagues over taught lessons.* Safety drivers' traits of learning difficulties with AI technologies and their lack of AI knowledge resemble the majority of end-users of AI systems [68]. Their cognitive preferences and learning characteristics during the training process can be extended to a broader range of lay users, informing the design of training strategies that are more aligned with user learning preferences to enhance users' understanding of AI technologies. Hence, we consider that when introducing AI systems to new users, the training programs are supposed to allow for active involvement, visible feedback, and quick response to ensure such systems can be fully understood and analyzed by the lay user. Some HCI researchers have explored novel interaction strategies to train users and improve their understanding of AI systems, such as combining theory

with practice through immersive technologies [51], collaborative learning assistant agents [49], interactive training games [27], etc. We hope the insights gained from this study regarding the technology learning and skill transfer characteristics of safety drivers can inform the design of AI system training strategies for a wider range of AI users.

***Mental model calibration and bidirectional communication.*** In our study, safety drivers constantly updated their mental models during their working practices. However, these changing mental models were not assessed and calibrated in a timely and effective manner. Our participants reported usually having misunderstandings about AV systems, as they lacked a clear understanding of the mechanisms of AV systems. They could only verify their mental models by comparing the results produced by AV systems. In addition, due to the lack of reliable calibration processes, they have to wait until the consequences aroused by their wrong mental model occur before they can calibrate their wrong mental model. Therefore, it is crucial to implement effective strategies that support users in calibrating their mental models timely through their day-to-day interactions with AI systems, for example providing real-time or regular feedback and clear and concise explanations about the mechanisms of the AI systems [61]. Our study also found that safety drivers struggled with communicating their understanding to the AV system, and the AV system was also unable to comprehend the safety drivers' mental models. This limited one-way communication may pose challenges to the partnerships between safety drivers and autonomous vehicles and can increase safety risks. Recent studies also showed that the ability to facilitate effective communication during interactions is more effective in promoting teamwork than technical capabilities [26, 44]. Aside from improving the interpretability and transparency of AV to users [31, 48, 73], AI systems should be able to evaluate the changes in users' mental models over time through building shared mental models and bidirectional communication between humans and machines [3, 64].

***Decision-making in high-risk situations.*** Trende's [80] and Adam's [57] studies investigated human decision-making in highly autonomous vehicles in time-critical situations, both of their results showed that subjects tended to accept the automated system's suggestions rather than decisions made by their own. Contrary to their studies, our study shows that in emergency and high-risk situations, safety drivers take over the AV more by instinct and intuition than by rational decision making. Most safety drivers report that they could not accurately and quickly obtain the real-time decision logic and results of AV, and in this case, they were more inclined to believe their own decisions. Safety drivers tend to rely more on instinct and intuition when taking over autonomous vehicles, as they are unable to accurately and quickly understand the rationale behind AV decision-making and make real-time evaluations. As a result, they prefer to trust their own decisions more. However, sometimes intrusive decision-making may bring risks. As Villemeur [83] mentioned, the likelihood of a human making an incorrect supervisory decision during a short time frame for unexpected, high-consequence scenarios is close to one hundred percent. We have analyzed the following questions to be solved:

*How to allow users to accurately assess the accuracy of AI decisions.* In this study, humans have the absolute power to interfere

in AI decisions, while wrong interventions may lead to serious consequences. Therefore, it is crucial to enable safety drivers to accurately assess the AI decisions and make a more informed and responsible control transfer.

*How to evaluate and predict which decision-making, either by humans or AI, is better for an event that did not occur.* Balance decision-making between AI systems and humans has remained a challenge in human-AI collaboration research. Researchers are actively exploring potential solutions, such as establish a third-party decision-making mechanism or evaluation mechanism between these two decision makers, which may be helpful to provide humans with evaluation information about both decision makers [4, 15].

*How to get humans to overcome their limitations (e.g., uncontrollable instinctive reactions, overconfidence) and willingly hand over control to AI, once the AI system is accurate in its decisions and humans have realized that AI is accurate.* Collaborative decision-making between humans and AI systems is becoming more widespread. Further exploration is needed to enable users to accurately understand the AI system and overcome any interaction barriers that may arise due to human limitations.

## 6.2 Towards Responsible AI: From Safety Drivers' Perspectives

**Support the upstream AV industry's "consequence awareness".** The creators of the upper echelons of technology often lack awareness of what the potential consequences of the technology they create may have [12]. This can lead to a reduction in the level of caution and responsibility among various stakeholders, which increases the likelihood of creating a "crazy machine" [77]. In this study, safety drivers, seen as the testers in the last piece, are required to drive AV on real roads to find defects. They have a better understanding of how the technology performs in the real world than upstream technology developers and are expected to identify accumulated defects and take responsibility for accumulated risks. It is crucial for stakeholders in the upstream autonomous vehicle industry to be aware of the potential consequences of the technology they are involved in. Organizations should also be responsible for embedding responsibility in all stages of technology production, especially those that are often neglected.

***Challenges in work practices.*** Every aspect of the AI industry is closely interconnected, from development, testing, to implementation. The bottom-tier AI workers still play a critical role in the AI industry, and many studies have highlighted the significant issues that arise from neglecting those workers [62, 70]. We found similar challenges in safety drivers' work practices.

It is paradoxical and worrying that safety drivers generally have a low knowledge of AV, but their job requires them to be able to predict AV accurately. According to our interviews, companies were more concerned with having safety drivers learn how to operate the machines and less concerned with their understanding of technological mechanisms. Often, safety drivers were only able to acquire AV knowledge through limited learning approaches, such as social media or informal communication with colleagues, and the knowledge acquired through these tracks is often incomplete and inaccurate. Lack of knowledge of technology hinders their work practices and can lead to misprediction [58], misattribution [89], etc., which may

lead to risks. We suggest that companies should value safety drivers' knowledge of AV and provide them with accessible learning sources.

Our findings also show that it can be challenging for safety drivers to transfer their training experience with autonomous vehicles to real-world road tests. The calibration of their knowledge and mental model of AV systems relies more on trial and error through hands-on practices than on training. However, this means they must bear real safety and violation risks in order to gain experience, which is a costly way to learn.

Although previous studies [22, 84] have explored how to help users gain AV operating experience, most of them are based on laboratory and simulated environments. There is limited research on how to effectively transfer this experience to the real world. Therefore, finding more effective training methods to reduce the cost of transferring skills for safety drivers and enable them to form accurate mental models without taking risks is a crucial challenge that needs to be addressed.

Additionally, it was reported that most AV companies focused more on safety drivers' initial training and neglected to provide adequate ongoing training and assessment during safety drivers' extended periods of work. It is recommended that organizations provide long-term mentorship for autonomous vehicle operators, so that they can effectively evaluate the performance and weaknesses of the system, especially as technology evolves.

***Involving front-line AI workers into human-centered AI research.*** Despite being the lowest-level testers in the AV industry and not directly involved in the optimization of AV, safety drivers have unique, first-hand experiences with autonomous vehicles as front-line workers who have been working closely with them for an extended period of time. They often develop their own understanding of AV flaws that may go unnoticed by the research and development team. Their opinions may assist the research and development team in identifying and correcting defects that have not yet triggered safety risks, and addressing them prior to the occurrence of risks to improve the safety of AV testing. However, according to our research, safety drivers, as marginal workers, do not have a strong voice and are frequently overlooked for the concerns they report. We recommend that the information feedback mechanism be improved and that the opinions and input of front-line automation workers be valued and taken into consideration. Moreover, since front-line workers like safety drivers have hands-on experience with AI, involving these workers in the technology development process would contribute to the advancement of the technology and lead to a more human-centered approach in AI research. Their perspectives and experiences could also be involved in the development, data collection, and analysis phases to help identify any potential issues or concerns with the AI system and ensure its usability and reliability.

## 6.3 Experience Migration: Spy on the Future User Experience of Automation Systems

As automation advancements progress, the integration of technology systems with human capabilities is creating novel driving experiences [87]. As the "first passenger of AV," safety drivers' experience of high-level autonomous driving can help us understand the

automation experience of a wider group. Through interviews, we learned about the positive and negative aspects of their experience, which were closely related to AV technologies. Additionally, we found that their social interactions with people in workplaces were facilitated by technologies, and exposure to emerging technologies increased their self-efficacy and desire for expression. We also observe a dynamic and complex change in the human-machine relationship, resulting from complex environments, different AI system qualities, and various personal preferences and personality traits. For example, some drivers felt a strong connection and externalized themselves to AV, whereas some felt weak ties to it. When facing external accusations caused by the AV, some empathized with the car, while some felt embarrassed and tried to cover up its deficiencies by takeover. Some even crossed the line by taking over the car without complying with company rules due to personal preferences or social pressure, and some chose to disassociate themselves with AV by gesturing or mouthing to other road participants when AV made mistakes. We learned their real-road driving experiences, which made them experience the pride of being noticed, unnecessary annoyances, and risks caused by outside curiosity and low acceptance. These safety drivers' experience can help us gain insight into the future user experience of broader AI technologies and automation systems.

## 7 LIMITATIONS

We acknowledge that our study is an initial investigation. We only interviewed a small sample size of safety drivers, and the sensitivity of their work may have deterred them from speaking more critically about their experiences, which may lead to subjective bias and a lack of generalizability for our results. To further work on this study, more information would be valuable and could be gathered by conducting more interviews with safety drivers and other stakeholders in the AV industry, for a more complete view. On that basis, more quantitative and confirmatory research could be conducted to validate the findings.

## 8 CONCLUSION

In this paper, we present how safety drivers perceive, understand, and partner with AV in the real world. We also found that, as front-line workers, safety drivers are forced to take risks accumulated from the AV industry upstream and are also confronting restricted self-development in working for AV development. We discussed the opportunities for human-vehicle partnerships and the improvement of workers' experiences. We compared our findings with previous literature and find the gaps between human-AV interactions in controlled experiment environment and real world long term practices. We contribute the first empirical evidence of the lived experience of safety drivers—the first passengers in the development of AV, as well as the grassroots workers for AV. We hope that this paper will provide valuable insights for more concerted and confirmatory research efforts in the area and consequently contribute toward the implementation of automated driving.

## ACKNOWLEDGMENTS

their lived experiences with us, as this study would not have been possible without them.

## REFERENCES

[1] Amirul Ibrahim Abu Bakar, Mohd Azman Abas, Mohd Farid Muhamad Said, and Tengku Azrul Tengku Azhar. 2022. Synthesis of Autonomous Vehicle Guideline for Public Road-Testing Sustainability. *Sustainability* 14, 3 (2022), 1456.

[2] Eugen Altendorf, Constanze Schreck, Gina Weßel, Yigiterkut Canpolat, and Frank Flemisch. 2019. Utility assessment in automated driving for cooperative human–machine systems. *Cognition, Technology & Work* 21, 4 (2019), 607–619.

[3] Robert W Andrews, J Mason Lilly, Divya Srivastava, and Karen M Feigh. 2022. The role of shared mental models in human-AI teams: a theoretical review. *Theoretical Issues in Ergonomics Science* (2022), 1–47.

[4] Verena Bader and Stephan Kaiser. 2019. Algorithmic decision-making? The user interface and its role for human involvement in decisions supported by artificial intelligence. *Organization* 26, 5 (2019), 655–672.

[5] Ronald M. Baecker. 2019. Automation, work, and jobs. *Computers and Society* (2019).

[6] SJ Baltrusch, F Krause, AW de Vries, W van Dijk, and MP de Looze. 2022. What about the Human in Human Robot Collaboration? A literature review on HRC's effects on aspects of job quality. *Ergonomics* 65, 5 (2022), 719–740.

[7] Marcel Baltzer, Eugen Altendorf, Sonja Meier, Frank Flemisch, N Stanton, S Landry, GD Bucchianico, and A Vallicelli. 2014. Mediating the interaction between human and automation during the arbitration processes in cooperative guidance and control of highly automated vehicles: basic concept and first study. *Advances in Human Aspects of Transportation Part I* (2014), 439–450.

[8] Victoria Banks, Emily Shaw, and David R Large. 2018. Keeping the driver in the loop: The 'Other'ethics of automation. In *Congress of the International Ergonomics Association*. Springer, 70–79.

[9] Mark J Brosnan. 2002. *Technophobia: The psychological impact of information technology*. Routledge.

[10] EunJeong Cheon, Cristina Zaga, Hee Rin Lee, Maria Luce Lupetti, Lynn Dombrowski, and Malte F Jung. 2021. Human-Machine Partnerships in the Future of Work: Exploring the Role of Emerging Technologies in Future Workplaces. In *Companion Publication of the 2021 Conference on Computer Supported Cooperative Work and Social Computing*. 323–326.

[11] Torkil Clemmensen and Torkil Clemmensen. 2021. Socio-Technical HCI Design in a Wider Context. *Human Work Interaction Design: A Platform for Theory and Action* (2021), 267–280.

[12] Mark Coeckelbergh. 2020. Artificial intelligence, responsibility attribution, and a relational justification of explainability. *Science and engineering ethics* 26, 4 (2020), 2051–2068.

[13] Mark Colley, Christian Bräuner, Mirjam Lanzer, Marcel Walch, Martin R. K. Baumann, and Enrico Rukzio. 2020. Effect of Visualization of Pedestrian Intention Recognition on Trust and Cognitive Load. *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (2020).

[14] SAE On-Road Automated Vehicle Standards Committee et al. 2014. Taxonomy and definitions for terms related to on-road motor vehicle automated driving systems. *SAE Standard J* 3016 (2014), 1–16.

[15] Laura Crompton. 2021. The decision-point-dilemma: Yet another problem of responsibility in human-AI interaction. *Journal of Responsible Technology* 7 (2021), 100013.

[16] Mitchell Cunningham and Michael A Regan. 2015. Autonomous vehicles: human factors issues and future research. In *Proceedings of the 2015 Australasian Road safety conference*, Vol. 14.

[17] Allan Dafoe, Yoram Bachrach, Gillian Hadfield, Eric Horvitz, Kate Larson, and Thore Graepel. 2021. Cooperative AI: machines must learn to find common ground.

[18] Mustafa Demir, Aaron D Likens, Nancy J Cooke, Polemnia G Amazeen, and Nathan J McNeese. 2018. Team coordination and effectiveness in human-autonomy teaming. *IEEE Transactions on Human-Machine Systems* 49, 2 (2018), 150–159.

[19] Nachiket Deo and Mohan M Trivedi. 2019. Looking at the driver/rider in autonomous vehicles to predict take-over readiness. *IEEE Transactions on Intelligent Vehicles* 5, 1 (2019), 41–52.

[20] Nachiket Deo and Mohan Manubhai Trivedi. 2020. Looking at the Driver/Rider in Autonomous Vehicles to Predict Take-Over Readiness. *IEEE Transactions on Intelligent Vehicles* 5 (2020), 41–52.

[21] Taşkın Dirsehan and Ceren Can. 2020. Examination of trust and sustainability concerns in autonomous vehicle adoption. *Technology in Society* 63 (2020), 101361.

[22] Vinayak V. Dixit, Sai Chand, and Divya Jayakumar Nair. 2016. Autonomous Vehicles: Disengagements, Accidents and Reaction Times. *PLoS ONE* 11 (2016).

[23] Mitch Downey. 2021. Partial automation and the technology-enabled deskilling of routine jobs. *Labour Economics* 69 (2021), 101973.

[24] Dániel A Drexler, Arpád Takács, Tamás D Nagy, and Tamás Haidegger. 2019. Handover Process of Autonomous Vehicles-technology and application challenges.

[25] *Acta Polytechnica Hungarica* 16, 9 (2019), 235–255.

[25] Yuemeng Du, Jingyan Qin, Shujing Zhang, Sha Cao, and Jinhua Dou. 2018. Voice User Interface Interaction Design Research Based on User Mental Model in Autonomous Vehicle. In *HCI*.

[26] Charles Duhigg. 2016. What Google learned from its quest to build the perfect team. *The New York Times Magazine* 26, 2016 (2016), 2016.

[27] Mahdi Ebnali, Cyrus Kian, Majid Ebnali-Heidari, and Adel Mazloumi. 2019. User experience in immersive VR-based serious game: an application in highly automated driving training. In *International Conference on Applied Human Factors and Ergonomics*. Springer, 133–144.

[28] Fredrick Ekman, Mikael Johansson, and Jana Sochor. 2016. Creating Appropriate Trust for Autonomous Vehicle Systems: A Framework for HMI Design.

[29] Oswald Mesumbe Ekwoge, Awdren Fontão, and Arilo C Dias-Neto. 2017. Tester experience: concept, issues and definition. In *2017 ieee 41st annual computer software and applications conference (compsac)*, Vol. 1. IEEE, 208–213.

[30] Madeleine Gibson, John Lee, Vindhya Venkatraman, Morgan Price, Jeffrey Lewis, Olivia Montgomery, Bilge Mutlu, Joshua Domeyer, and James Foley. 2016. Situation awareness, scenarios, and secondary tasks: measuring driver performance and safety margins in highly automated vehicles. *SAE International Journal of Passenger Cars-Electronic and Electrical Systems* 9, 1 (2016), 237–243.

[31] Balint Gyevnar, Massimiliano Tamborski, Cheng-Hsien Wang, Christopher G. Lucas, Shay B. Cohen, and Stefano V. Albrecht. 2022. A Human-Centric Method for Generating Causal Explanations in Natural Language for Autonomous Vehicle Motion Planning. *ArXiv* abs/2206.08783 (2022).

[32] Michelle Hester, Kevin Lee, and Brian P. Dyre. 2017. "Driver Take Over": A Preliminary Exploration of Driver Trust and Performance in Autonomous Vehicles. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 61 (2017), 1969 – 1973.

[33] Charles P. Hewitt, Ioannis Politis, Theocharis Amanatidis, and Advait Sarkar. 2019. Assessing public perception of self-driving cars: the autonomous vehicle acceptance model. *Proceedings of the 24th International Conference on Intelligent User Interfaces* (2019).

[34] Gaojian Huang and Brandon Pitts. 2020. Age-related differences in takeover request modality preferences and attention allocation during semi-autonomous driving. In *International Conference on Human-Computer Interaction*. Springer, 135–146.

[35] Michael Huesemann and Joyce Huesemann. 2011. *Techno-fix: why technology won't save us or the environment*. New Society Publishers.

[36] Nidhi Kalra and Susan M. Paddock. 2016. Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability? *Transportation Research Part A-policy and Practice* 94 (2016), 182–193.

[37] Hannu Karvonen, Iina Aaltonen, Mikael Wahlström, Leena Salo, Paula Savioja, and Leena Norros. 2011. Hidden roles of the train driver: A challenge for metro automation. *Interacting with computers* 23, 4 (2011), 289–298.

[38] Hyung Jun Kim and Ji Hyun Yang. 2017. Takeover requests in simulated partially autonomous vehicles considering human factors. *IEEE Transactions on Human-Machine Systems* 47, 5 (2017), 735–740.

[39] Jungsook Kim, Hyunsuk Kim, Woojin Kim, and Daesub Yoon. 2018. Take-over performance analysis depending on the drivers' non-driving secondary tasks in automated vehicles. *2018 International Conference on Information and Communication Technology Convergence (ICTC)* (2018), 1364–1366.

[40] Anirudh Kishore Bhoopalam, Roy van den Berg, Niels Agatz, and Caspar Chorus. 2021. The long road to automated trucking: Insights from driver focus groups. *Roy and Agatz, Niels AH and Chorus, Caspar, The Long Road to Automated Trucking: Insights from Driver Focus Groups (February 4, 2021)* (2021).

[41] Jeamin Koo, Jungsuk Kwac, Wendy Ju, Martin Steinert, Larry J. Leifer, and Clifford Nass. 2015. Why did my car just do that? Explaining semi-autonomous driving actions to improve driver understanding, trust, and performance. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 9 (2015), 269–275.

[42] Jiin-in Lee, Na-eun Kim, and Jin-woo Kim. 2017. A study on driver experience in autonomous car based on trust and distrust model of automation system. *Journal of Digital Contents Society* 18, 4 (2017), 713–722.

[43] Y. Li, Dihua Sun, Min Zhao, Dong Chen, Senlin Cheng, and Fei Xie. 2018. Switched Cooperative Driving Model towards Human Vehicle Copiloting Situation: A Cyberphysical Perspective. *Journal of Advanced Transportation* 2018 (2018), 1–11.

[44] Claire Liang, Julia Proft, Erik Andersen, and Ross A Knepper. 2019. Implicit communication of actionable information in human-ai teams. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13.

[45] Patrick Lindemann, Tae-Young Lee, and Gerhard Rigoll. 2018. Catch My Drift: Elevating Situation Awareness for Highly Automated Driving with an Explanatory Windshield Display User Interface. *Multimodal Technol. Interact.* 2 (2018), 71.

[46] Robert E Llaneras, Brad R Cannon, and Charles A Green. 2017. Strategies to assist drivers in remaining attentive while under partially automated driving: Verification of human–machine interface concepts. *Transportation research record* 2663, 1 (2017), 20–26.

[47] Duri Long and Brian Magerko. 2020. What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors*

*in computing systems.* 1–16.

[48] Joseph B. Lyons. 2013. Being Transparent about Transparency: A Model for Human-Robot Interaction. In *AAAI Spring Symposium: Trust and Autonomous Systems*.

[49] Ioannis Magnisalis, Stavros Demetriadis, and Anastasios Karakostas. 2011. Adaptive and intelligent systems for collaborative learning support: A review of the field. *IEEE transactions on Learning Technologies* 4, 1 (2011), 5–20.

[50] Moira Maguire and Brid Delahunt. 2017. Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Ireland Journal of Higher Education* 9, 3 (2017).

[51] Sotiris Makris, Panagiotis Karagiannis, Spyridon Koukas, and Aleksandros-Stereos Matthaiakis. 2016. Augmented reality system for operator support in human–robot collaborative assembly. *CIRP Annals* 65, 1 (2016), 61–64.

[52] James Manyika, Susan Lund, Michael Chui, Jacques Bughin, Jonathan Woetzel, Parul Batra, Ryan Ko, and Saurabh Sanghvi. 2017. Jobs lost, jobs gained: Workforce transitions in a time of automation. *McKinsey Global Institute* 150 (2017).

[53] Nora McDonald, Sarita Schoenebeck, and Andrea Forte. 2019. Reliability and inter-rater reliability in qualitative research: Norms and guidelines for CSCW and HCI practice. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–23.

[54] Daniel V McGehee, Mark Brewer, Chris Schwarz, Bryant Walker Smith, et al. 2016. *Review of automated vehicle technology: policy and implementation implications.* Technical Report. Iowa. Dept. of Transportation.

[55] Seamus McGuinness, John Will Freebairn, and Kostas G Mavromaras. 2008. *Characteristics of minimum wage employees.* Australian Fair Pay Commission.

[56] Raissa Pokam Meguia, Serge Debernard, Christine Chauvin, and Sabine Langlois. 2019. Principles of transparency for autonomous vehicles: first results of an experiment with an augmented reality human–machine interface. *Cognition, Technology & Work* (2019), 1–14.

[57] Adam Millard-Ball. 2016. Pedestrians, Autonomous Vehicles, and Cities. *Journal of Planning Education and Research* 38 (2016), 12 – 6.

[58] Alexander G Mirnig, Philipp Wintersberger, Christine Sutter, and Jürgen Ziegler. 2016. A framework for analyzing and calibrating trust in automated vehicles. In *Adjunct proceedings of the 8th international conference on automotive user interfaces and interactive vehicular applications.* 33–38.

[59] Brian Mok, Mishel Johns, Key Jung Lee, David Sirkin, Page Ive, and Wendy Ju. 2015. Emergency, automation off: Unstructured transition timing for distracted drivers of automated vehicles. In *2015 IEEE 18th international conference on intelligent transportation systems.* IEEE, 2458–2464.

[60] António Moniz and Bettina-Johanna Krings. 2014. Technology assessment approach to human-robot interactions in work environments. In *2014 7th International Conference on Human System Interactions (HSI).* IEEE, 282–289.

[61] Shane T Mueller, Robert R Hoffman, William Clancey, Abigail Emrey, and Gary Klein. 2019. Explanation in human-AI systems: A literature meta-review, synopsis of key ideas and publications, and bibliography for explainable AI. *arXiv preprint arXiv:1902.01876* (2019).

[62] Michael Muller, Christine T Wolf, Josh Andres, Michael Desmond, Narendra Nath Joshi, Zahra Ashktorab, Aabhas Sharma, Kristina Brimijoin, Qian Pan, Evelyn Duesterwald, et al. 2021. Designing ground truth and the social life of labels. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–16.

[63] Lisa Orii, Diana Tosca, Andrew L Kun, and Orit Shaer. 2021. Perceptions on the Future of Automation in r/Truckers. In *Extended abstracts of the 2021 CHI conference on human factors in computing systems.* 1–6.

[64] Scott Ososky, David Schuster, Florian Jentsch, Stephen Fiore, Randall Shumaker, Christian Lebiere, Unmesh Kurup, Jean Oh, and Anthony Stentz. 2012. The importance of shared mental models and shared situation awareness for transforming robots from tools to teammates. In *Unmanned systems technology XIV*, Vol. 8387. SPIE, 397–408.

[65] Lawrence A Palinkas, Sarah M Horwitz, Carla A Green, Jennifer P Wisdom, Naihua Duan, and Kimberly Hoagwood. 2015. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and policy in mental health and mental health services research* 42, 5 (2015), 533–544.

[66] Ilias E. Panagiotopoulos, George J. Dimitrakopoulos, Gabriele Keraite, and Urte Steikuniene. 2020. Are Consumers Ready to Adopt Highly Automated Passenger Vehicles? Results from a Cross-national Survey in Europe. In *VEHITS*.

[67] Sarvapali D Ramchurn, Sebastian Stein, and Nicholas R Jennings. 2021. Trustworthy human-AI partnerships. *Iscience* 24, 8 (2021), 102891.

[68] Mireia Ribera and Agata Lapedriza. 2019. Can we do better explanations? A proposal of user-centered explainable AI.. In *IUI Workshops*, Vol. 2327. 38.

[69] Herbert J Rubin and Irene S Rubin. 2011. *Qualitative interviewing: The art of hearing data.* sage.

[70] Nithya Sambasivan, Shivani Kapania, Hannah Highfill, Diana Akrong, Praveen Paritosh, and Lora M Aroyo. 2021. "Everyone wants to do the model work, not the data work": Data Cascades in High-Stakes AI. In *proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–15.

[71] Beau G Schelble, Christopher Flathmann, Nathan J McNeese, Guo Freeman, and Rohit Mallick. 2022. Let's Think Together! Assessing Shared Mental Models, Performance, and Trust in Human-Agent Teams. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–29.

[72] Shervin Shahrdar, Luiza Menezes, and Mehrdad Nojoumian. 2018. A survey on trust in autonomous systems. In *Science and Information Conference.* Springer, 368–386.

[73] Yuan Shen, Shanduojiao Jiang, Yanlin Chen, Eileen Jianxun Yang, Xilun Jin, Yuliang Fan, and Katherine Driggs Campbell. 2020. To Explain or Not to Explain: A Study on the Necessity of Explanations for Autonomous Vehicles. *ArXiv* abs/2006.11684 (2020).

[74] R Jay Shively, Joel Lachter, Summer L Brandt, Michael Matessa, Vernol Battiste, and Walter W Johnson. 2017. Why human-autonomy teaming?. In *International conference on applied human factors and ergonomics.* Springer, 3–11.

[75] Nasim Sonboli, Jessie J Smith, Florencia Cabral Berenfus, Robin Burke, and Casey Fiesler. 2021. Fairness and transparency in recommendation: The users' perspective. In *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization.* 274–279.

[76] Eric Strobl and Robert J Thornton. 2002. Do large employers pay more in developing countries? The case of five African countries. *The Case of Five African Countries (December 2002)* (2002).

[77] Cass Robert Sunstein. 2018. Algorithms, Correcting Biases. *Artificial Intelligence - Law* (2018).

[78] Árpád Takács, Dániel András Drexler, Péter Galambos, Imre J Rudas, and Tamás Haidegger. 2018. Assessment and standardization of autonomous vehicles. In *2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES).* IEEE, 000185–000192.

[79] Duy Tran, Jianhao Du, Weihua Sheng, Denis Osipychev, Yuge Sun, and He Bai. 2018. A human-vehicle collaborative driving framework for driver assistance. *IEEE Transactions on Intelligent Transportation Systems* 20, 9 (2018), 3470–3485.

[80] Alexander Trende, Anirudh Unni, Lars Weber, Jochem W. Rieger, and Andreas Lüdtke. 2019. An investigation into human-autonomous vs. human-human vehicle interaction in time-critical situations. *Proceedings of the 12th ACM International Conference on PErvasive Technologies Related to Assistive Environments* (2019).

[81] Paola Tubaro, Antonio A Casilli, and Marion Coville. 2020. The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. *Big Data & Society* 7, 1 (2020), 2053951720919776.

[82] Christopher J Turner, Ruidong Ma, Jingyu Chen, and John Oyekan. 2021. Human in the Loop: Industry 4.0 technologies and scenarios for worker mediation of automated manufacturing. *IEEE access* 9 (2021), 103950–103966.

[83] Alain Villemeur. 1992. Assessment, hardware, software and human factors, volume 2 of Reliability, availability, maintainability and safety assessment.

[84] Marcel Walch, Tobias Sieber, Philipp Hock, Martin R. K. Baumann, and Michael Weber. 2016. Towards Cooperative Driving: Involving the Driver in an Autonomous Vehicle's Decision Making. *Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (2016).

[85] Ding Wang, Shantanu Prabhat, and Nithya Sambasivan. 2022. Whose AI Dream? In search of the aspiration in data annotation.. In *CHI Conference on Human Factors in Computing Systems.* 1–16.

[86] Jun Wang, Li Zhang, Yanjun Huang, and Jian Zhao. 2020. Safety of autonomous vehicles. *Journal of advanced transportation* 2020 (2020).

[87] Gesa Wiegand, Kai Holländer, Katharina Rupp, and Heinrich Hussmann. 2020. The Joy of Collaborating with Highly Automated Vehicles. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications.* 223–232.

[88] Gesa Wiegand, Matthias Schmidmaier, Thomas Weber, Yuanting Liu, and Heinrich Hussmann. 2019. I Drive - You Trust: Explaining Driving Behavior Of Autonomous Cars. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (2019).

[89] David D Woods, Leila J Johannesen, Richard I Cook, and Nadine B Sarter. 1994. *Behind human error: Cognitive systems, computers and hindsight.* Technical Report. Dayton Univ Research Inst (Urdi) OH.

[90] Yang Xing, Chen Lv, Dongpu Cao, and Peng Hang. 2021. Toward human-vehicle collaboration: Review and perspectives on human-centered collaborative automated driving. *Transportation Research Part C: Emerging Technologies* (2021).

[91] Wei Xu. 2020. From automation to autonomy and autonomous vehicles: Challenges and opportunities for human-computer interaction. *Interactions* 28, 1 (2020), 48–53.

[92] Wei Xu, Marvin J Dainoff, Liezhong Ge, and Zaifeng Gao. 2023. Transitioning to human interaction with AI systems: New challenges and opportunities for HCI professionals to enable human-centered AI. *International Journal of Human–Computer Interaction* 39, 3 (2023), 494–518.

[93] Shiyan Yang, Steven E Shladover, Xiao-Yun Lu, John Spring, David Nelson, and Hani Ramezani. 2018. A first investigation of truck drivers' on-the-road experience using cooperative adaptive cruise control. (2018).

[94] Yucheng Yang, Burak Karakaya, Giancarlo Caccia Dominioni, Kyosuke Kawabe, and Klaus Bengler. 2018. An HMI Concept to Improve Driver's Visual Behavior and Situation Awareness in Automated Vehicle. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (2018), 650–655.

[95] Yucheng Yang, Burak Karakaya, Giancarlo Caccia Dominioni, Kyosuke Kawabe, and Klaus Bengler. 2018. An hmi concept to improve driver's visual behavior and situation awareness in automated vehicle. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 650–655.

[96] Jiehuang Zhang, Ying Shu, and Han Yu. 2021. Human-Machine Interaction for Autonomous Vehicles: A Review. In *International Conference on Human-Computer Interaction*. Springer, 190–201.