Chapter 2-Empirical Research for Mixed Traffic Research

Dengbo HE
https://orcid.org/0000-0003-4359-4083
Intelligent Transportation Thrust
Robotics and Autonomous Systems Thrust
Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, Guangdong, China

Xiaoyu ZHOU
https://orcid.org/0009-0000-2214-5647
Robotics and Autonomous Systems Thrust
Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, Guangdong, China

Yang LIU
https://orcid.org/0009-0000-1416-0265
Intelligent Transportation Thrust
Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, Guangdong, China

Jiahao ZHANG
https://orcid.org/0009-0002-3833-3911
Robotics and Autonomous Systems Thrust
Hong Kong University of Science and Technology (Guangzhou)
Guangzhou, Guangdong, China

Abstract

This chapter focuses on data collection methods for empirical data collection in the mixed traffic domain, including driving simulation research, field studies, naturalistic driving research, and observational research. We will discuss the pros and cons of the methods and the basic steps for data preparation. Examples of the research with the above-mentioned approaches in the mixed traffic domain will also be discussed.

2.1 Why Empirical Research

The survey-based approach has been widely adopted for mixed traffic research in the past few years and has shown a great advantage in understanding users' concerns and acceptance of autonomous driving before the technology matures. For example, T. Zhang et al., (2020) explored the factors influencing users' acceptance of autonomous vehicles (AVs) through a survey study including 647 Chinese drivers. Similarly, based on a survey study with 451 valid samples, Wu et al., (2023) found that the level of automation can directly or indirectly (through trust) affect public acceptance of AV while demographic characteristics of respondents can only indirectly affect public acceptance of AV through trust. However, the survey data can hardly facilitate an understanding of the micro and meso behaviors for several reasons. First, up until now, fully autonomous vehicles (AVs) have not been officially commercialized and a few people have experienced or even witnessed fully AVs running on public roads. Thus, their attitudes towards AVs were mostly based on their limited knowledge about AVs and thus their attitudes might be biased in such situations (Janatabadi & Ermagun, 2022). Second, data-driven, learning-based, or even model-based research all rely on users' micro and meso behaviors in mixed traffic, which cannot be captured in survey studies. Thus, collecting empirical data in mixed traffic is vital for the optimization of AV control algorithms and mixed traffic management. In this Chapter, we will mainly review the research that adopted empirical approaches, i.e., an approach that gains knowledge or obtains empirical evidence through direct and indirect observation or experience. Specifically, we will discuss four commonly used approaches in driving behavioral research, including driving simulation research, field studies, naturalistic driving research, and observational research. We will briefly discuss the pros and cons of the approaches, introduce the basic procedures for conducting research with these approaches, and discuss the findings generated from research adopting these research methods.

2.2 Types of Empirical Data Collection Methods Adopted for Mixed Traffic Research

A. Driving Simulation Approach

(1) Driving Simulator Studies

The driving simulation is the most widely adopted approach in the driving behavior research domain. In driving simulator studies for mixed traffic, experimenters can design the traffic scenarios and have participants act as either drivers or other road agents (e.g., pedestrians or cyclists) in the scenarios. Given that AVs can be easily simulated in the simulator, drivers' responses to the AVs in mixed traffic can be readily captured at a relatively low cost in a well-controlled environment. For example, with the driving simulator, all participants in a study can experience the same scenarios and the experimenters can easily control the order of the experimental conditions, which is especially beneficial for isolating the impacts from the factors of interest. The highly risky scenarios or tasks can also be tested in driving simulators. However, the pros of the driving simulator may also bring cons. For example, to isolate the influential factors in a driving simulator experiment, the complexity and thus the reality of the scenarios might be compromised. The lack of real risk and adverse consequences in driving simulator experiments may also bias users' behaviors and raise questions on the validity of the experiment outcomes, especially when quantitative data is needed.

(2) Types of Driving Simulators

The validity of the driving simulator studies highly depends on the fidelity of the simulator. Though no strict standards define the level of fidelity of a driving simulator and previous research usually framed their simulators as low-, medium-, and high-fidelity ones. In general, this categorization is based on the hardware of the system, including the degrees of freedom (DoF), the cabin fidelity, and the display systems. For example, a computer system without motion functions (i.e., fixed-base) and non-cabin setup is usually considered a low-fidelity one

(Figure 2.2.1a). The fixed system, or the system with low DoF (i.e., ≤3) and with a cabin-like setup can be considered as a medium one (Figure 2.2.1b). While the high-fidelity simulators usually have a high DoF (≥6) motion platform, over 180-degree view angle, and at least a quarter-cabin (Figure 2.2.1c). However, the fidelity, by definition, should be more than hardware-related and can also be related to the simulation software. For example, in a train simulator study, (Olsson, 2023) found that the driving performance of train drivers in physically low-fidelity but functionally high-fidelity simulators is comparable to real train driving performance. Functional fidelity, however, is rarely mentioned in the driving research, potentially because functions related to the specific research purposes usually work well, and hence can be regarded as with high fidelity in most cases; while the functions non-related to the specific research are either well-controlled (e.g., some driving simulators cannot turn the steering when lane centering control works) or not being used in the research. Table 2.2.1 summarizes the description of the simulators in some of the mixed-traffic-related research.

Table 2.2.1 Descriptions of the simulators used in mixed-traffic-related research up to 2023

Research	Description of the simulator	Self-identified level of fidelity
Manual drivers' experience and driving behavior in repeated interactions with automated Level 3 vehicles in mixed traffic on the highway (Stange et al., 2022a)	"The driving simulation was programmed using SILAB 6.0 and consists of a physical mock-up seat box with driver and passenger seats. The three LCD projectors (1920 × 1080 pixels each) projected the scenery onto an array of three screens (2 m × 2 m each) covering a field of view of 180°"	Medium-fidelity simulator
Safety at first sight? – Manual drivers' experience and driving behavior at first contact with Level 3 vehicles in mixed traffic on the highway (Stange et al., 2022b)		
The impact of expectations about automated and manual vehicles on drivers' behavior: Insights from a mixed traffic driving simulator study (Miller et al., 2022)	"Consisted of a vehicle mockup with an adjustable seat, FanaTech steering wheel, and pedals with force feedback…Run with SILAB 6.5 on an RTX3070 PC and three 55-inch 4K displays with a field of view of about 160 degrees"	Not mentioned
Driver-automated vehicle interaction in mixed traffic: Types of interaction and drivers' driving Styles (Ma & Zhang, 2024)	"Each simulator consisted of a vehicle mockup with an adjustable seat, FanaTech steering wheel, and pedals with force feedback ... The simulation was run with SILAB 6.5 on an RTX3070 PC and three 55-inch 4K displays with a field of view of about 160 degrees."	Not mentioned
Effects of marking automated vehicles on human drivers on highways (Fuest et al., 2020)	"The basis of the static driving simulator was a BMW 6 series mockup. A 6-channel projection system provided	Medium-fidelity simulator

	a realistic driving environment, with a refresh rate of 60 Hz. Three projectors were used for the 180° front view, and three projectors for the rear view (side and rear mirrors). We used the driving simulation software SILAB 6.5 of the Würzburg Institute for Traffic Sciences GmbH [29] and logged the driving data with 240 Hz. A 6-channel noise simulation completed the driving simulation. A freely programmable instrument cluster was used as human-machine interface. A tachometer and a speedometer were implemented for displaying driving-relevant information in this study."	
Risk modeling and quantification of a platoon in mixed traffic based on the mass-spring-damper model (Jiang et al., 2020)	"mainly includes a Logitech G29 vehicle controller (steering wheel, pedal, and gear lever) and two sorts of simulation software (PreScan and Matlab/Simulink). "	Not mentioned
Driver behavior at a freeway merge to mixed traffic of conventional and connected autonomous vehicles (Chityala et al., 2020)	"DriveSafety Model CDS-250 Driving Simulato. the drivers sit in a partial cab based on a Ford Focus sedan. the simulator consists of various computing units, with most on the automotive frame containing the sedan, and a separate station for authoring scenarios for the simulations."	Not mentioned
Effect of cognitive distraction on physiological measures and driving performance in traditional and mixed traffic environments (Hua et al., 2021)	"The driving environment was established using the simulation software UC-win/Road and displayed on three 32-in LED displays. The horizontal viewing angle of the scenario display system was 120°, The sampling frequency was 100 Hz."	Low-fidelity simulator
Learning in mixed traffic: Drivers' adaptation to ambiguous communication depending on their expectations toward automated and manual vehicles (Miller et al., 2023)	"a vehicle mockup featuring an adjustable seat, force-feedback racing wheel (FanaTec Base V2), and pedals (Fanatec CSL Elite). 17-inch touchscreen mounted in the position of the vehicle's center console. vehicle's front view was simulated on three 55-inch 4K UHD LED TVs, which were placed in front of the mockup, and created a driver's view	Medium-fidelity simulator

	of around 120 degrees.”	
Cooperative driving in mixed traffic of manned and unmanned vehicles based on human driving behavior understanding (J. Lu et al., 2023)	“The driving simulator has 3 connected monitors for in-vehicle view, an extra monitor for control and data display, and a Logitech G290 driving force suit (a steering wheel, pedals, and a shifter).”	Not mentioned
Development of a research testbed for cooperative driving in mixed traffic of human-driven and autonomous vehicles (J. Lu et al., 2022)		
The effect of visual advanced driver assistance systems on a following human driver in a mixed-traffic condition (Ariansyah et al., 2023)	“The simulator consisted of three screens combined providing 36 degrees and 175 degrees on the vertical and horizontal field of view (FoV), respectively. It was also equipped with the steering wheel, pedals, and adjustable driver seat, while the control of the vehicle was implemented with automatic transmission.”	Low-fidelity simulator

(a) (b) (c)

Figure 2.2.1. Driving simulation systems at the Hong Kong University of Science and Technology (Guangzhou): a) low-fidelity simulator; b) medium-fidelity simulator; c) high-fidelity simulator.

In addition to research with a single driving simulator, in recent years, to better understand users’ behaviors in mixed traffic, researchers also started to conduct research with multiple driving simulators, or the combination of driving simulators and walking simulators. For example, Kalantari et al., (2023) replicated vehicle-pedestrian interaction in a safe and controllable virtual environment (32 pairs, one driver and one pedestrian interacting with each other in different scenarios). The results show that kinematic cues play a greater role than psychological characteristics such as sensory seeking and social value orientation in determining who passes an unmarked intersection first. In Miller et al., (2022), multi-agent simulators were also used to investigate how the role of expectations shapes drivers’ reactions to AVs and human-driven vehicles and found that human driving behavior can be improved when AVs match one’s expectations. These types of research can provide insights into the more complex gaming processes in human-AV or AV-pedestrian interactions.

B. Field Studies and On-Road Studies

Given that the relatively low validity of the driving simulation might be a concern for specific research questions, field studies can be an alternative option, which has relatively low risk (as compared to experiments on open public roads) and high validity but still reserves some levels of experiment controllability. In general, the field studies are conducted in closed tracks or test fields without public traffic. In the experiment, participants are usually required to conduct specific tasks in instrumented vehicles provided by the experimenters. Different sensors are installed onto the instrumented vehicles, for example, eye tracking devices, physiological sensors, radars, and on-board vehicle data collection devices. A. Zhou et al., (2023) collected eye-tracking data to investigate different Connected and Autonomous Vehicle (CAV) control settings on users' acceptance of CAV and their distraction behaviors. However, it should be noted that, as the experiment was still conducted in a contrived environment, the participants' behaviors may still be skewed to some extent. Specifically, the participants would be accompanied by the experimenters and irrelevant traffic would be removed. Hence, the participants would experience lower than usual risks.

To simulate the risk more realistically, researchers may also conduct on-road studies using instrumented vehicles. Being different from field studies, the participants may drive the vehicles equipped with data collection devices on public roads, while accompanied by experimenters. As can be imagined, the fidelity of the scenarios will be higher compared to field studies and participants may experience higher levels of risk on public roads. However, with such an approach, the controllability of the scenarios will the lower compared to the field experiments, as although the experimenters can control some static factors related to the scenario (e.g., the time of day, the weather, and the experimental areas), they have little to no control over the traffic scenarios surrounding the instrument vehicles.

At the same time, given that fully AVs are not mature enough, to understand drivers' or pedestrians' responses to the AVs, the Wizard-of-Oz (Dahlbäck et al., 1993) technique was commonly adopted. With this approach, the system was simulated (e.g., by having an experimenter driving a non-autonomous vehicle) but the participants were not informed of such setup. Hence, participants may perceive the vehicle as fully autonomous. With this approach, previous researchers have successfully simulated the operation of AVs on the road. For example, Detjen et al., (2020) used the Wizard-of-Oz method to simulate an autonomous driving system. All participants were told that the vehicle was controlled by the driving automation, but it was actually controlled by an experimenter. Users' feedback and behavioral data were collected. Given that some (4 out of 12) participants perceived the system as human-operated, the authors compared how the trust in the AV varied across the two groups of participants and it was found that participants trusted the human driver more.

C. Naturalistic Driving Research

The naturalistic driving study is a powerful tool for obtaining realistic driving behaviors among drivers. Being different from field studies, sensors, and equipment are installed in a vehicle that can run on public roads. The vehicle can either be the participants' own vehicle (Liu et al.,

2023) or the vehicle that is rented to the participants (e.g., Advanced Vehicle Technology Consortium, (Hong et al., 2021)). With such a setup, in naturalistic driving studies, the experimenters can observe realistic driving behaviors with minimum interference from the experimenters or the context of the research. However, the disadvantage is also obvious. Given that the study is usually conducted on public roads, the experimenters had little to no control over the traffic scenarios and states of the drivers. Thus, the density of the meaningful data points for a specific research question might be sparse and data cleaning would be time-consuming. This type of research also has high requirements for data storage and transfer. For example, the largest and the most commonly used naturalistic driving dataset, the SHRP2 contains over 2 PB data for nearly 2360 participants, leading to 3700 participant-years data in total (Hallmark et al., 2015). The following table summarizes some of the existing public datasets from naturalistic driving studies.

Table 2.2.2 Available public naturalistic driving studies up to 2023

Dataset	Vehicle type	Total length of data duration	Locations of data collection	Types of collected data	Research
UDRIVE	Cars, trucks and powered two-wheelers	88,000 hours of vehicle data	Six European Unions Member States	Video & Vehicle Sensors Data	(Eenink et al., 2014)
CNDS	Conventional vehicle	Over 53,000 vehicle hours	Canada	Video of Driver, Video & Vehicle Sensors Data	(Klauer et al., 2018)
ANDS	Private vehicle	360 volunteer drivers for 4 months	NSW and Victoria of Australian	Video of driver & other road users' behaviour, video & sensors data of vehicle	(Williamson et al., 2015)
IVBSS	Light vehicles	Over 213,000 miles accumulated	Parts of southeast Michigan, US	Video & Sensors Data of Drivers & Vehicles; Assessment of Drivers	(Steve, 2011)
SHNDS	Cars	Over 5,000 hours of continuous driving data from 60 drivers	Shanghai city of China	Video of driver, video & sensors data of vehicle, assessment	(Guo et al., 2022)
NTDS	Private vehicle	nearly 102,000 trips over 800,000 km (500,000 miles).	Commonwealth of Virginia in the United States	Video of driver, video & sensors data of vehicle	(Lee et al., 2011)

SHRP2	Conventional vehicle	3700 participant-years	Several cities in US	Video & Vehicle Sensors Data	(National Research Council (U.S.), 2001)

Another major limitation of the naturalistic driving approach when applied to mixed traffic research is that, so far, existing naturalistic driving research has all focused on drivers' behaviors in the ego-vehicle. Although some research has used vehicles with some levels of driving automation (Wen et al., 2023), they provide limited knowledge on how human drivers respond to other AVs or vehicles controlled by driving automation in mixed traffic. Such kind of research is limited by two constraints at this stage. On the one hand, to collect the interactions with other road agents on the road, the ego-vehicle will need to be equipped with appropriate external sensors (e.g., LiDAR and cameras), which are expensive and may lead to privacy concerns. On the other hand, even when the instrumented vehicles with necessary sensors are available, collecting data regarding mixed traffic is still difficult, given that only a small portion of vehicles are equipped with advanced driving assistant systems (ADASs) or autonomous driving systems. However, it should be pointed out that knowing how human drivers respond to automated road agents and evaluate users' cognitive and decision-making processes would facilitate more efficient and explainable AV control algorithms. With more AVs being available on public roads, future naturalistic driving research may better capture human-AV interactions from an ego-vehicle-driver behavioral perspective of view.

D. Observational Approach

Though capturing human-AV interactions through a naturalistic driving research approach is difficult due to the sparsity of the AVs on public roads at this stage, observational studies can be a feasible option at this stage and have been widely adopted in previous research, given that certain cities or areas have allowed the AVs to test and operate commercially on public roads. Depending on the resolution of the datasets, the observational datasets can be categorized into two types: the trajectory-oriented one and the safety-oriented one. The former focuses more on the micro trajectories of human-AV interactions on public roads while the latter focuses more on the safety records of mixed traffic.

In traditional traffic research, the trajectory-oriented data can be collected through roadside units (e.g., (U.S. Department of Transportation Federal Highway Administration, n.d.)), on-road data collection vehicles (e.g., (K. Chen et al., 2024)), and unmanned aerial vehicles (e.g., (Krajewski et al., 2018)). The Next Generation Simulation (NGSIM) Open Data used cameras mounted on highways to collect vehicle trajectories on highways. However, given the sparsity of the AVs on the road, collecting mixed traffic data with fixed sensors or cameras is inefficient at this stage. Thus, they may not be suitable for mixed traffic research until AVs start to saturate the market.

To date, existing trajectory-oriented mixed traffic datasets were mostly collected using AV-mounted sensors. The three most representative ones are the Waymo Open Dataset, the Lyft Level-5 Dataset, and the nuScenes dataset. All of them provided perceived speed and location information of the road agents surrounding the AVs at a relatively high sampling frequency (e.g., 10Hz in Waymo Open Dataset). Thanks to the perception capabilities of the AVs, the datasets contain information regarding the road agents that are hundreds of meters away from the AV. Thus, representative traffic scenarios such as car-following behaviors and lane-changing behaviors can be extracted. Specifically, the Waymo Open Dataset consists of two parts: perception and motion. The perception part contains 1,000 20-second video clips, each of which is composed of well-synchronized and calibrated high-resolution LiDAR and camera data recorded in urban and suburban areas. The motion part consists of 103,354 20-second video clips representing 574 hours of driving data collected over 1,750 km of roadways. The Lyft Level-5 Dataset contains information collected by 7 cameras and 3 lidars on a fleet of 25 SAE Level-5 AVs operated by Lyft, which includes map information covering more than 4,000 roads, 197 crosswalks, 60 stop signs, and 54 parking areas. Visible traffic participants, including vehicles, pedestrians, and cyclists within each scene were also detected, and their motions such as velocity, acceleration, and yaw rate were provided. In total, the dataset also includes more than 55,000 3D human-annotated frames. The nuScenes dataset was collected by autonomous vehicles equipped with LiDARs, radars, and cameras in four cities around the world and it contains 1200 hours of driving.

To date, many studies have been conducted with the Waymo Open Dataset and the Lyft Level-5 dataset. For example, Wen et al explored the strategies human drivers took when following AVs versus human-driven vehicles based on the trajectory data from the Waymo dataset. Similarly, Li et al extracted human-following AV and human-following human events in the Lyft Level-5 dataset. However, it should be noted that the original objectives of the three datasets were for AV control algorithm development and were not designed for mixed-traffic behavioral research. For certain reasons, all these datasets chose to cut the video clips into 20-second-long discrete segments and the information other than the location of the surrounding road agents has been removed. As a result, these datasets are not suitable for long-duration driver behavior extraction or complex scenarios that may exceed 20 seconds. Similarly, due to the sparsity of the crashes on the road, the trajectory-oriented datasets are not suitable for research that aims to understand the impact of AVs on traffic safety and analyze the patterns of AV-involved crashes. Thus, local authorities and governments have also initiated plans to collect information regarding AV-involved crashes. The two most widely used datasets include the NHTSA dataset (by the National Highway Traffic Safety Administration) and the CA-DMV dataset (by the California Department of Motor Vehicle), both involved data regarding AV-involved crashes. Detailed information on these two datasets and other available datasets are listed in Table 2.2.3.

Table 2.2.3 Available public crash/trajectory datasets involving driving automation.

Datasets	Data collection duration	Types of involved road agents	Locations of data collection	Total length of data	Type of data	Related research
CA DMV	2014-2019	Conventional & High-Level Autonomous Vehicles	California	712 autonomous vehicle collision reports received	AV collision & disengagement records	(Sinha et al., 2021)
NHTSA AV crash data archive	2008-2009	Conventional & High-Level Autonomous Vehicles	Several cities in US	Over 1100 autonomous vehicle crash records	Accidents records	(National Highway Traffic Safety Administration, 2013)
AV crash dataset	2014-2024	Conventional & Autonomous Vehicles (with ADAS)	Several cities in US	Total of 2236 pieces of data	Crash data from CA DMV, NHTSA, news & land use, weather, and geometry information	(Transport research center, 2024)
NGSIM	2016	Conventional vehicle	Emeryville, Los Angeles, and Atlanta	2 hours 30 minutes	Video Data	(U.S. Department of Transportation Federal Highway Administration, n.d.)
HighD	2018	Conventional vehicle	German Highways	147 hours	Video Data	(Krajewski et al., 2018)
Lyft Level 5	2019	Autonomous vehicle	Palo Alto, California	1,000 hours	Video & Radar Data	(Kesten et al., n.d.)
Waymo	2021-2024	Autonomous vehicle	25 Cities in US	10 hours 50 minutes	Video & Radar Data	(K. Chen et al., 2024)
nuScenes	2018-2023	Autonomous vehicle	Boston and Singapore	5 hours 33 minutes	Video & Radar Data	(motional, 2023)

However, although the data from the observational studies can maintain the highest level of validity, it is not omnipotent. Specifically, given that the researchers usually have no access to the road users involved in certain traffic events, it is usually unlikely to collect information regarding the road users' psychological states or processes. Hence, the data from this type of research can hardly support the analysis of psychological or cognitive factors of human behaviors (e.g., how users may allocate their attention differently when interacting with AVs). Further, as the data was reported by the AV fleet operators, drivers, or police, the quality of the data might be compromised, and minor events (without leading to large property or injuries) may be under-reported.

E. Choice of Approaches for Mixed Traffic Research

In general, no single research method so far can support research on all types of research questions in mixed traffic. It is vital to understand the pros and cons of the research approaches when they are used for specific research questions. In addition to the monetary and time costs and constraints of the available resources, we may select the research approaches by considering the factors in the following three dimensions.

First, the validity of the data for specific research questions. For example, given that the data from the driving simulators have lower ecological validity compared to the data from the on-road studies, the data collected from on-road studies may be more suitable for developing quantitative models to inform the design of the AV algorithms.

Second, the availability of the data. Not all approaches can provide all the necessary information for all research questions. For example, although observational studies may provide more realistic data, the collection of demographic information and psychological states of road users would be difficult. While naturalistic driving studies can capture both realistic driving behaviors and demographic information of the drivers, the human-AV interaction samples might be too sparse at this stage, given that only a few regions or states have allowed the AVs to be operated on public roads. Hence, simulator studies or field experiments may better reveal how psychological factors may be associated with drivers' behaviors in mixed traffic.

Third, the controllability of the experimental conditions. For example, in both naturalistic and observational studies, experimenters can hardly control the traffic scenarios. Hence, large datasets are needed to extract rare events that may meet their criteria. In this case, the simulator or field study approach can be a better option to obtain data in a relatively short period of time, given that a higher level of experiment control can be guaranteed. Experimental control is also vital to identify the influential factors of certain behaviors, as consistent scenarios can reduce the variations uncounted by the factors of interest.

Regarding the research approaches that can be used for mixed traffic research, we can summarize their pros and cons in these three dimensions in Table 2.2.4. It should be noted that the pros and cons of the research methods may vary with the progress of time. For example, when more AVs are operating on public roads, it is possible to collect enough data with naturalistic driving studies.

Table 2.2.4 Comparisons of the research approaches for mixed traffic research.

Approaches	Data validity	Data availability	Experimental control
Driving simulation	Low to medium	Depends on the needs of the research questions	High
Field study	Medium		Medium to high
On-road study	Medium to high		Low to medium
Naturalistic driving research	High		Low
Observational approach	High		Low

2.3 Data Type and Data Processing in Empirical Research

One of the basic questions one may ask when selecting the approaches is the types of data that can be collected with the approach. In general, for mixed traffic research, the following types of data were used in previous research, i.e., trajectory data, behavioral data, subjective data, and eye-tracking data.

A. Trajectory Data

(1) Application of the Trajectory Data

The trajectory data can capture the dynamic processes of interactions among road agents in mixed traffic. In general, previous research used the trajectory data in two ways. The raw and cleaned trajectory data was more commonly used for research that focused on AV algorithm designs (e.g., motion planning). For example, Gu et al., (2020) used Waymo trajectory data to design a car following an algorithm based on a modified Long-Short-Term-Memory (LSTM). At the same time, some other studies relied more on the metrics extracted from the trajectory data. For example, in Wen et al., (2023), the headway distance and vehicle speed extracted from the trajectories of the leading and following vehicles were used to model the car-following behaviors of human drivers in mixed traffic. In another study, again, using the Waymo dataset, T. Li et al., (2023) analyzed and compared how drivers responded to human-driven vehicles versus AVs differently, by modeling the effect of road agent types (AV or human-driven vehicles) on drivers' selection of headways in car-following and overtaking events. The trajectory data can also be used to evaluate the safety of mixed traffic. Given the sparsity of the AV-involved crashes, previous research also extracted surrogate safety metrics from the trajectory data. For example, Q.-L. Lu et al., (2021) compared the performance of different algorithms for collision prediction in mixed traffic based on the estimated TTC in the lane-changing event.

(2) Data Processing for Trajectory Data

As mentioned previously, the trajectory data was mostly collected on the road (either public or closed-track), for example, from field studies (Huang et al., 2020), on-road studies (Y. Chen et al., 2020), naturalistic studies (G. Li et al., 2023), and observational studies (Basu & Saha, 2022). Thus, the data may contain lots of noise and if used inappropriately, may lead to inaccurate model training or behavioral predictions in follow-up studies (M. Zhou et al., 2017). Hence, the raw trajectory data needs to be processed before being used. The main purpose of data processing for trajectory data is as follows.

- **Data format conversion**: Data collected by different sensors may have different data formats, so they need to be converted into a unified format for subsequent data processing and analysis. Special techniques may be required to obtain the necessary data. For example, computer visions (e.g., image resizing (Khajeh Hosseini et al., 2022), color space conversion (Baskaran et al., 2017) and feature extraction (Samanta et al., 2018)) may be required to extract road agents and identify their locations and movements from the video and image data. Similarly, pre-processing is also needed for the LiDAR data before trajectories can be extracted (Llorca et al., 2010; Yan et al., 2011).
- **Data annotation**: The data collected by some sensors may need to be annotated for training and testing purposes. For example, for image or video data, the objects and road signs may need to be annotated manually or semi-automatically before the trajectory can be extracted; for LIDAR data, it is also necessary to annotate the position and size of the objects. Yan et al., (2011) explores how to semantically annotate different types of trajectory data, this includes annotating image, video or LiDAR data for training and testing purposes. It introduces a framework or methodology to semantically annotate these heterogeneous trajectory data in order to provide a basis for subsequent data processing and analysis. Llorca et al., (2010) deals with methods for the collection and processing of traffic data in V2I networks. This includes the annotation of sensor data to facilitate the enhancement and analysis of floating vehicle data. C. Chen et al., (2020) proposed a framework for representing and compressing vehicle trajectory data, in which the data collected by the sensors will be labeled to facilitate the efficient representation and compression of vehicle trajectory data.
- **Outlier detection and handling**: Abnormal situations (e.g., sensor failure or extreme environmental conditions) may lead to errors or strong noise in the collected data. In such cases, outlier detection and processing of data may be required to eliminate the strong influence of the outliers. Typical outlier detection methods include:
  - Z-score method (Hodge & Austin, 2004): This method identifies outliers by calculating the z-score of each data point and flagging those that fall outside a certain threshold.
  - Interquartile range (IQR) method (Yang et al., 2019): This method involves calculating the IQR of the data and identifying outliers as those that fall below the first quartile minus 1.5 times the IQR or above the third quartile plus 1.5 times the IQR.
  - Modified z-score method (Sandbhor & Chaphalkar, 2019): Similar to the z-score method, but it is more robust to outliers and works better with skewed distributions.
  - Mahalanobis distance (Ghorbani, 2019): This method calculates the distance of each data point from the centroid of the data distribution and flags those that fall beyond a certain threshold.
  - Local outlier factor (LOF) (Breunig et al., 2000): This method compares the density of data points in their local neighborhood to identify outliers.
- **Data imputation**: In data acquisition and transmission processes, some sensor data may be missing or removed due to errors. Thus, interpolation and other methods may be used to impute missing values. Typical interpolation methods include:
  - Linear interpolation (Kuffel et al., 1997): This method estimates missing values by drawing a straight line between the two nearest known data points and using the position along that line to estimate the missing value.

- Polynomial interpolation (de Boor & Ron, 1990): This method fits a polynomial function to the known data points and uses it to estimate missing values.
- Spline interpolation (De Boor, 1968): This method uses a piecewise polynomial function to interpolate between data points, providing a smooth estimate for missing values.
- Kriging (van Beers & Kleijnen, 2004): This method is commonly used in geostatistics and spatial analysis to estimate unknown values based on the spatial correlation between known data points.
- Moving average interpolation (Thompson, 1947): This method replaces missing values with the average of neighboring data points, which can help to smooth out fluctuations in the data.

B. Behavioral Data

(1) Application of the Behavioral Data

Although the trajectory data can reveal the interactions among the road agents, it may only provide superficial information regarding the outcome of road agents' decisions, but not how and why human road agents adopted the strategy. As a direct measure of how human road agents respond to AVs or other human-operated road agents on the road, the behavioral data may provide more insights into the impact of AVs on traffic safety and efficiency. For example, when investigating how pedestrians may interact with AVs, in addition to deep learning approaches that focused on trajectory prediction (Bhujel & Yau, 2023), researchers also focused on the pedestrians' hesitation time and body movements when encountering an AV (Rodríguez Palmeiro et al., 2018). In combination with questionnaire data, theoretical models regarding the relationships between the psychological states and demographic features of road users and their behaviors could be established. For example, Papadimitriou et al., (2016) delves into the impact of various psychological variables on pedestrian decision-making and interactions with traffic signals. This study can provide valuable insights into understanding how psychological factors influence pedestrians' behavior in urban environments, which can guide future endeavors to guide road users in mixed traffic.

(2) Data Processing for Behavioral Data

Similar to the trajectory data, the behavioral data can be highly noisy. The first step in extracting behavioral data is the clear definition of the specific behaviors. Usually, this includes the pre-conditions for an action (e.g., only after the AV reaches 50 meters, a leg movement can be considered as a street crossing intention, (Pillai, 2017)), the definition of the action (e.g., the movement of any leg or the forward-leaning of the upper body), and the time frame that an action can be counted (e.g., only the actions happened between the first glance to the cue and the event onset can be regarded as a pre-event action (He et al., 2021)).

Though we would try to clearly define the behaviors of interest, subjectivity inevitably has to be involved in the definition of certain behaviors. For example, in He et al., (2021), pre-event actions can be defined as any movement that is deemed as preparation for an upcoming event. Though the pre-conditions and the time frame can be strictly defined, the actions are subject to raters' judgment. Thus, to reduce the subjectivity in the behavior extraction, multiple raters would be involved. Preliminary extraction would be conducted to refine the criteria for behaviors, followed by discussions to reach a consensus and then independent judges by each rater.

However, still, given that the judgment of the behavioral data from each rater can be subjective, a common standard for evaluating the quantification of the behavior identification is inter-rater reliability (IRB). Typical IRB metrics include Cohen's Kappa (Ihejirika et al., 2015) and the intra-class correlation coefficient (Bobak et al., 2018). If the IRB metrics are lower than acceptable, then a refinement of the criteria is needed, and a re-extraction of behaviors based on the new standard may need to be conducted unless the metrics reach the threshold.

C. Subjective Data

(1) Application of the Subjective Data

With behavioral data, although we can look into one's decisions in specific scenarios, it is still difficult to answer why the behavior differs and how we can shape one's behaviors. Further, the behavioral data can still hardly inform one's cognitive procedures leading to actions. Thus, to better model road users' behaviors, subjective data can be collected. For example, with the trajectory data only, we can rebuild their behavioral patterns, but we can hardly answer why human drivers took different strategies when following AVs versus following human-driven vehicles. Thus, it is difficult for us to design countermeasures to guide road users' behaviors in mixed traffic. To resolve this issue, a previous study (Zhao et al., 2020) rebuilt car-following scenarios in closed tracks; in addition, they also evaluated users' trust in AVs using a subjective questionnaire. With such an approach, the relationship between a psychological state (i.e., trust) and the car-following behaviors can be established.

However, so far, a gap between the psychological findings and the data-driven approach still exists. More specifically, although we can identify how subjective metrics can affect road users' behaviors, or even build theoretical models explaining the variance accounted by some psychological factors, these qualitative findings are not well utilized in the data-driven approach and can hardly guide the design of the AV control algorithms. Future research is still needed to bridge the gap between the psychology-based approach and the trajectory-/behavior-related approach.

(2) Data Processing for Subjective Data

In most cases, the subjective data can be collected through questionnaires, interviews, and focus groups. The subjective data can be collected before, during, and after an experiment. Given that subjective data is related to what the participants feel or subjectively perceive the scenarios, the validity and reliability of the data might be questioned and need to be evaluated, especially if a self-designed questionnaire is used. Here, the validity reflects whether the subjective measure can accurately measure what it is supposed to measure, which includes three dimensions (i.e., construct validity, content validity, and criterion validity); while the reliability evaluates to what extent the measure can collect the same data in repeated observations, which also has three dimensions (i.e., test-retest reliability, interrater reliability, and internal reliability). Some typical metrics regarding validity and reliability include Cronbach α (Tavakol & Dennick, 2011) and KMO (Prato et al., 2005). The readers could refer to the reference for more information regarding evaluating the validity and reliability. For the design of the questionnaire, interview, or focus group, the readers could refer to (Gubrium & Holstein, 2002; Stanton et al., 2017) for more information.

D. Physiological Data and Eye-Tracking Data

(1) Application of the Physiological Data and Eye-tracking Data

As the subjective data is sensitive to users' bias and the validity can sometimes be questionable, more objective data that can reflect road users' psychological states is sometimes preferred. It is widely acknowledged that the changes in some psychological states can lead to variations in physiological indices. For example, stress is associated with variations in the galvanic skin responses (Fernandes et al., 2014), heart rate, and heart rate variability (Kim et al., 2018; Taelman et al., 2009). Hence, previous research also adopted physiological measures to evaluate users' acceptance of external human-machine interface (eHMI) in mixed traffic (e.g., (Tan et al., 2022)). The electroencephalogram (EEG) has also been widely used in previous research. By monitoring the electrical activity of the brain, researchers can infer a driver's or pedestrian's emotional state and cognitive load when they are faced with stress, emergencies, or changes in comfort, which can further be used to infer road users' capability to respond to traffic incidents (P. Li et al., 2022; Luque et al., 2024). Physiological signals can also be combined with other sensor data for comprehensive analysis (Rao et al., 2023).

As for the eye-tracking measures, it is estimated that around 90% of needed information in driving is visual information (Sivak, 1998). Thus, understanding road users' road behaviors can provide insights into their attention allocation strategies. Eye trackers are widely adopted to collect road users' visual behaviors. With eye trackers, experimenters can identify the gaze movements of human road agents and reveal their attention allocation and perception patterns, thus enabling a deep understanding of their behaviors. For example, eye-tracking technologies have been widely used in pedestrian-AV interaction studies to understand how the eHMI design can influence the pedestrian's perception of the AVs and hence behaviors when encountering AVs. In addition, the shared attention between human road users and AVs may facilitate more efficient and safe interactions among them, given that the computer vision domain has already emphasized the importance of directing attention to areas of importance in traffic scenarios (e.g.,(Makrigiorgos et al., 2019)).

(2) Data Processing for Physiological Data and Eye-tracking Data

The physiological data is subjective to noises in the environment. Thus, filtering noise is usually required as the first step in processing physiological data. For example, the electrocardiogram (ECG) may have some baseline drift, which would downgrade the detection of the R peaks in the signal. Thus, as common practice, the raw signal would be detrended before being fed into the following steps. The EEG signal is also sensitive to the artifacts and the background noise in the environment. Thus, in addition to collecting data in the magnetic shielding room, the bandpass filter is usually used to filter out the utility frequency (e.g., 50 Hz in China and

60 Hz in the U.S.) in the EEG signals. After pre-processing, depending on the way the signals are used, specific metrics can be extracted from the raw signal, or the cleaned signals can be fed into some algorithms directly. For example, the inter-beat-intervals (iBi) can be extracted from the ECG signal, and be further processed into the heart rate and the heart rate variability (HRV). The HRV can then be transformed into the frequency domain and metrics such as the LF (power of the low-frequency) and HF (power of the power of the how-frequency) can be extracted. Alternatively, the raw ECG data can be fed into the algorithms directly without extracting handcrafted metrics; this is especially common in the deep learning domain so that the models might be able to learn the high-level features in the raw data (Xu et al., 2019).

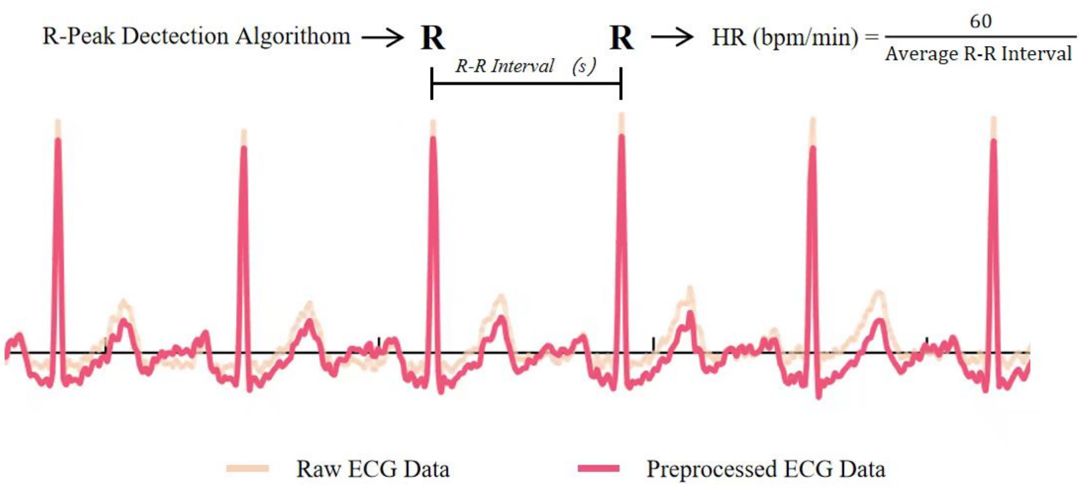

Figure 2.3.1 Typical ECG data.

The eye-tracking data include two categories, gaze-related data and non-gaze-related data. As can be inferred from the name, the gaze-related data is about the location where the road agents are looking. This category of data is directly related to users' attention allocation. A number of metrics can be extracted from the gaze-related data, including but not limited to glance duration, fixation duration, frequency of the glances, percentage of time looking at a specific area of interest, and spatial density. For more information regarding the definition of gaze-related data in the driving domain, the readers could refer to the ISO standards (e.g., ISO_PRF_15007-1_(E) and ISO_TS_15007-2_(E)). At the same time, the non-gaze-related data also contains rich information regarding users' states. For example, the percentage time of eye closure (PERCLOS) has been found to be closely related to one's fatigue levels (J. Zhang et al., 2021); and the size of the pupil is sensitive to the workload one is experiencing (Klingner et al., 2008).

2.4 Experiment Design Basics for Empirical Research

After deciding the types of data to be collected, another key task in conducting empirical studies is the experiment design. In general, when designing an empirical study, the experimenters need to consider three aspects of techniques, i.e., how to **measure** the variables of interest, how to **manipulate** some property of an actor-behavior-context (i.e., making a variable have a particular predetermined value or level for certain scenarios to be studied, and other specific values or levels for certain other scenarios so that the effect of the variable can be assessed by comparing those two sets of scenarios), and how to **control** the impact of variables that are important but irrelevant in a particular study.

In general, the experiment design can be regarded as balancing the pros and cons of different techniques. For example, to understand the influence of two eHMI designs on pedestrians' street crossing behavior, we can either let one participant experience multiple eHMIs (a within-subject design, see Table 2.4a), or have multiple participants experiencing different eHMIs respectively (a between-subjects design, see Table 2.4b). If a between-subjects design is adopted, then we may need to worry if there are individual differences between participants in different groups that may shadow the effect of eHMI so that the matching technique will need to be considered (i.e., making sure specific features of the participants across groups to be similar). However, in some scenarios, we may not know which features would influence their behaviors and thus it is difficult to match the participants across groups. Alternatively, if a within-subject design is adopted, we may need to worry about the learning effect and the fatigue effect (i.e., participants might learn to perform better or become tired with the progress of the experiment), if multiple trials were conducted for each participant. Thus, the counterbalancing technique will need to be used. However, this will increase the number of required participants, as a full counterbalance would significantly increase the required number of participants (e.g., a minimum of 6 participants for 3 levels, 24 orders for 4 levels, and 120 orders for 5 levels). In summary, the experiment design is more than an art rather than a technique, and that is also why multiple empirical studies may be needed for researchers to draw relatively solid conclusions on a research question. For more information regarding the experiment design and the following data analysis, especially the choice of the statistical models, the readers could refer to other books (e.g., Experimental Design: Procedures for the Behavioral Sciences by (Maydeu-Olivares & Millsap, 2009)).

Reference List

Ariansyah, D., Pardamean, B., & Caruso, G. (2023). The effect of visual advanced driver assistance systems on a following human driver in a mixed-traffic condition. Procedia Computer Science, 216, 221–229. https://doi.org/10.1016/j.procs.2022.12.130

Baskaran, S., Fang, S., & Jiang, S. (2017). Spatiotemporal visualization of traffic paths using color space time curve. 2017 IEEE International Conference on Big Data (Big Data), 3398–3405. https://doi.org/10.1109/BigData.2017.8258325

Basu, S., & Saha, P. (2022). Evaluation of risk factors for road accidents under mixed traffic: Case study on Indian highways. IATSS Research, 46(4), 559–573. https://doi.org/10.1016/j.iatssr.2022.09.004

Bhujel, N., & Yau, W.-Y. (2023). Disentangling Crowd Interactions for Pedestrians Trajectory Prediction. IEEE Robotics and Automation Letters, 8(5), 3078–3085. https://doi.org/10.1109/LRA.2023.3262420

Bobak, C. A., Barr, P. J., & O'Malley, A. J. (2018). Estimation of an inter-rater intra-class correlation coefficient that overcomes common assumption violations in the assessment of health measurement scales. BMC Medical Research Methodology, 18(1), 93. https://doi.org/10.1186/s12874-018-0550-6

Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: Identifying density-based local outliers. Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, 93–104. https://doi.org/10.1145/342009.335388

Chen, C., Ding, Y., Guo, S., & Wang, Y. (2020). DAVT: An Error-Bounded Vehicle Trajectory Data Representation and Compression Framework. IEEE Transactions on Vehicular Technology, 69(10), 10606–10618. https://doi.org/10.1109/TVT.2020.3015214

Chen, K., Ge, R., Qiu, H., AI-Rfou, R., Qi, C. R., Zhou, X., Yang, Z., Ettinger, S., Sun, P., Leng, Z., Baniodeh, M., Bogun, I., Wang, W., Tan, M., & Anguelov, D. (2024). WOMD-LiDAR: Raw Sensor Dataset Benchmark for Motion Forecasting (arXiv:2304.03834). arXiv. https://doi.org/10.48550/arXiv.2304.03834

Chen, Y., Chen, S., Xiao, T., Zhang, S., Hou, Q., & Zheng, N. (2020). Mixed Test Environment-based Vehicle-in-the-loop Validation—A New Testing Approach for Autonomous Vehicles. 2020 IEEE Intelligent Vehicles Symposium (IV), 1283–1289. https://doi.org/10.1109/IV47402.2020.9304658

Chityala, S., Sobanjo, J. O., Erman Ozguven, E., Sando, T., & Twumasi-Boakye, R. (2020). Driver Behavior at a Freeway Merge to Mixed Traffic of Conventional and Connected Autonomous Vehicles. Transportation Research Record: Journal of the Transportation Research Board, 2674(11). https://trid.trb.org/View/1739354

Dahlbäck, N., Jönsson, A., & Ahrenberg, L. (1993). Wizard of Oz studies—Why and how. Knowledge-Based Systems, 6(4), 258–266. https://doi.org/10.1016/0950-7051(93)90017-N

De Boor, C. (1968). On the convergence of odd-degree spline interpolation. Journal of Approximation Theory, 1(4), 452–463. https://doi.org/10.1016/0021-9045(68)90033-6

de Boor, C., & Ron, A. (1990). On multivariate polynomial interpolation. Constructive Approximation, 6(3), 287–302. https://doi.org/10.1007/BF01890412

Detjen, H., Pfleging, B., & Schneegass, S. (2020). A Wizard of Oz Field Study to Understand Non-Driving-Related Activities, Trust, and Acceptance of Automated Vehicles. 12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 19–29. https://doi.org/10.1145/3409120.3410662

Eenink, R., Barnard, Y., Baumann, M., Augros, X., & Utesch, F. (2014, April). UDRIVE: The European naturalistic driving study [Proceedings Paper]. TRA 2014. Proceedings of

Transport Research Arena; IFSTTAR. http://tra2014.traconference.eu/papers/pdfs/TRA2014_Fpaper_19523.pdf

Fernandes, A., Helawar, R., Lokesh, R., Tari, T., & Shahapurkar, A. V. (2014). Determination of stress using Blood Pressure and Galvanic Skin Response. 2014 International Conference on Communication and Network Technologies, 165–168. https://doi.org/10.1109/CNT.2014.7062747

Fuest, T., Feierle, A., Schmidt, E., & Bengler, K. (2020). Effects of Marking Automated Vehicles on Human Drivers on Highways. Information, 11(6), 286. https://doi.org/10.3390/info11060286

Ghorbani, H. (2019). MAHALANOBIS DISTANCE AND ITS APPLICATION FOR DETECTING MULTIVARIATE OUTLIERS. Facta Universitatis, Series: Mathematics and Informatics, 0, Article 0. https://doi.org/10.22190/FUMI1903583G

Gu, Z., Li, Z., Di, X., & Shi, R. (2020). An LSTM-Based Autonomous Driving Model Using a Waymo Open Dataset. Applied Sciences, 10(6), Article 6. https://doi.org/10.3390/app10062046

Gubrium, J. F., & Holstein, J. A. (2002). Handbook of Interview Research: Context and Method. SAGE.

Guo, F., Han, S., & Hankey, J. (2022). The Shanghai Naturalistic Driving Study (SHNDS).

Hallmark, S. L., Tyner, S., Oneyear, N., Carney, C., & McGehee, D. (2015). Evaluation of driving behavior on rural 2-lane curves using the SHRP 2 naturalistic driving study data. Journal of Safety Research, 54, 17–27. https://doi.org/10.1016/j.jsr.2015.06.017

He, D., Kanaan, D., & Donmez, B. (2021). In-vehicle displays to support driver anticipation of traffic conflicts in automated vehicles. Accident Analysis & Prevention, 149, 105842. https://doi.org/10.1016/j.aap.2020.105842

Hodge, V., & Austin, J. (2004). A Survey of Outlier Detection Methodologies. Artificial Intelligence Review, 22(2), 85–126. https://doi.org/10.1023/B:AIRE.0000045502.10941.a9

Hong, Y., Min, J., King, C. B., & Meeker, W. Q. (2021). Reliability Analysis of Artificial Intelligence Systems Using Recurrent Events Data from Autonomous Vehicles (arXiv:2102.01740). arXiv. http://arxiv.org/abs/2102.01740

Hua, Q., Jin, L., Jiang, Y., Guo, B., & Xie, X. (2021). Effect of Cognitive Distraction on Physiological Measures and Driving Performance in Traditional and Mixed Traffic Environments. Journal of Advanced Transportation, 2021, 1–17. https://doi.org/10.1155/2021/6739071

Huang, Y., Ding, H., Zhang, Y., Wang, H., Cao, D., Xu, N., & Hu, C. (2020). A Motion Planning and Tracking Framework for Autonomous Vehicles Based on Artificial Potential Field Elaborated Resistance Network Approach. IEEE Transactions on Industrial Electronics, 67(2), 1376–1386. https://doi.org/10.1109/TIE.2019.2898599

Ihejirika, R. C., Thakore, R. V., Sathiyakumar, V., Ehrenfeld, J. M., Obremskey, W. T., & Sethi, M. K. (2015). An assessment of the inter-rater reliability of the ASA physical status score in the orthopaedic trauma population. Injury, 46(4), 542–546. https://doi.org/10.1016/j.injury.2014.02.039

Janatabadi, F., & Ermagun, A. (2022). Empirical evidence of bias in public acceptance of autonomous vehicles. Transportation Research Part F: Traffic Psychology and Behaviour, 84, 330–347. https://doi.org/10.1016/j.trf.2021.12.005

Jiang, L., Ji, J., Ren, Y., Wang, H., & Huang, Y. (2020). Risk Modeling and Quantification of a Platoon in Mixed Traffic Based on the Mass-Spring-Damper Model. Journal of Advanced Transportation, 2020(Article ID 7475682). https://trid.trb.org/View/1722395

Kalantari, A. H., Yang, Y., Garcia de Pedro, J., Lee, Y. M., Horrobin, A., Solernou, A., Holmes, C., Merat, N., & Markkula, G. (2023). Who goes first? A distributed simulator stud y of vehicle-pedestrian interaction. Accident; Analysis and Prevention, 186, 107050. https://doi.org/10.1016/j.aap.2023.107050

Kesten, R., Usman, M., Houston, J., & Pandya, T. (n.d.). Lyft Level 5 AV Dataset 2019. Retrieved June 15, 2024, from http://archive.org/details/lyftlevel5avdataset2019

Khajeh Hosseini, M., Talebpour, A., Devunuri, S., & Hamdar, S. H. (2022). An unsupervised learning framework for detecting adaptive cruise control operated vehicles in a vehicle trajectory data. Expert Systems with Applications, 208, 118060. https://doi.org/10.1016/j.eswa.2022.118060

Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y. H., & Koo, B.-H. (2018). Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. Psychiatry Investigation, 15(3), 235–245. https://doi.org/10.30773/pi.2017.08.17

Klauer, C., Pearson, J., & Hankey, J. (2018). An Overview of the Canada Naturalistic Driving and Canada Truck Naturalistic Driving Studies. 7th International Symposium on Naturalistic Driving Research Inn at Virginia Tech.

Klingner, J., Kumar, R., & Hanrahan, P. (2008). Measuring the task-evoked pupillary response with a remote eye tracker. Proceedings of the 2008 Symposium on Eye Tracking Research & Applications, 69–72. https://doi.org/10.1145/1344471.1344489

Krajewski, R., Bock, J., Kloeker, L., & Eckstein, L. (2018). The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. 2018 21st International Conference on Intelligent Transportation Systems (ITSC), 2118–2125. https://doi.org/10.1109/ITSC.2018.8569552

Kuffel, P., Kent, K., & Irwin, G. (1997). The implementation and effectiveness of linear interpolation within digital simulation. International Journal of Electrical Power & Energy Systems, 19(4), 221–227. https://doi.org/10.1016/S0142-0615(96)00045-2

Lee, S. E., Simons-Morton, B. G., Klauer, S. E., Ouimet, M. C., & Dingus, T. A. (2011). Naturalistic assessment of novice teenage crash experience. Accident Analysis & Prevention, 43(4), 1472–1479. https://doi.org/10.1016/j.aap.2011.02.026

Li, G., Jiao, Y., Knoop, V. L., Calvert, S. C., & Van Lint, J. W. C. (2023). Large Car-Following Data Based on Lyft Level-5 Open Dataset: Following Autonomous Vehicles vs. Human-Driven Vehicles. 2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC), 5818–5823. https://doi.org/10.1109/ITSC57777.2023.10422574

Li, P., Li, Y., Yao, Y., Wu, C., Nie, B., & Li, S. E. (2022). Sensitivity of Electrodermal Activity Features for Driver Arousal Measurement in Cognitive Load: The Application in Automated Driving Systems. IEEE Transactions on Intelligent Transportation Systems, 23(9), 14954–14967. https://doi.org/10.1109/TITS.2021.3135266

Li, T., Han, X., Ma, J., Ramos, M., & Lee, C. (2023). Operational safety of automated and human driving in mixed traffic environments: A perspective of car-following behavior. Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability, 237(2), 355–366. https://doi.org/10.1177/1748006X211050696

Liu, Y., Abediasl, H., Ansari, A., & Shahbakhti, M. (2023). Characterizing driving behavior and link to fuel consumption for university campus shuttle minibuses. Canadian Society for Mechanical Engineering (CSME) / Computational Fluid Dynamics Society of Canada (CFD Canada). International Congress (2023 : Université de Sherbrooke, Québec), 6, ID 212. https://savoirs.usherbrooke.ca/handle/11143/21171

Llorca, D. F., Sotelo, M. A., Sánchez, S., Ocaña, M., Rodríguez-Ascariz, J. M., & García-Garrido, M. A. (2010). Traffic Data Collection for Floating Car Data Enhancement in V2I Networks. EURASIP Journal on Advances in Signal Processing, 2010(1), 719294. https://doi.org/10.1155/2010/719294

Lu, J., Hossain, S., Sheng, W., & Bai, H. (2023). Cooperative Driving in Mixed Traffic of Manned and Unmanned Vehicles based on Human Driving Behavior Understanding. 2023 IEEE International Conference on Robotics and Automation (ICRA), 3532–3538. https://doi.org/10.1109/ICRA48891.2023.10160282

Lu, J., Stracener, R., Sheng, W., Bai, H., & Hossain, S. (2022). Development of a Research Testbed for Cooperative Driving in Mixed Traffic of Human-driven and Autonomous Vehicles. 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 11403–11408. https://doi.org/10.1109/IROS47612.2022.9981436

Lu, Q.-L., Yang, K., & Antoniou, C. (2021). Crash risk analysis for the mixed traffic flow with human-driven and connected and autonomous vehicles. 2021 IEEE International Intelligent Transportation Systems Conference (ITSC), 1233–1238. https://doi.org/10.1109/ITSC48978.2021.9564897

Luque, F., Armada, V., Piovano, L., Jurado-Barba, R., & Santamaría, A. (2024). Understanding Pedestrian Cognition Workload in Traffic Environments Using Virtual Reality and Electroencephalography. Electronics, 13(8), Article 8. https://doi.org/10.3390/electronics13081453

Ma, Z., & Zhang, Y. (2024). Driver-Automated Vehicle Interaction in Mixed Traffic: Types of Interaction and Drivers' Driving Styles. Human Factors: The Journal of the Human Factors and Ergonomics Society, 66(2), 544–561. https://doi.org/10.1177/00187208221088358

Makrigiorgos, A., Shafti, A., Harston, A., Gerard, J., & Faisal, A. A. (2019). Human Visual Attention Prediction Boosts Learning & Performance of Autonomous Driving Agents (arXiv:1909.05003). arXiv. https://doi.org/10.48550/arXiv.1909.05003

Maydeu-Olivares, A., & Millsap, R. E. (2009). The SAGE Handbook of Quantitative Methods in Psychology. 1–800.

Miller, L., Koniakowsky, I. M., Kraus, J., & Baumann, M. (2022). The Impact of Expectations about Automated and Manual Vehicles on Drivers' Behavior: Insights from a Mixed Traffic Driving Simulator Study. Proceedings of the 14th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 150–161. https://doi.org/10.1145/3543174.3546837

Miller, L., Kraus, J., Koniakowsky, I., Pichen, J., & Baumann, M. (2023). Learning in Mixed Traffic: Drivers' Adaptation to Ambiguous Communication Depending on Their Expectations toward Automated and Manual Vehicles. International Journal of Human–Computer Interaction, 39(16), 3268–3287. https://doi.org/10.1080/10447318.2023.2215097

motional. (2023). nuScenes. Nuscenes. https://www.nuscenes.org/

National Highway Traffic Safety Administration. (2013). Motor Vehicle Traffic Crashes as a Leading Cause of Death in the United States, 2008 and 2009. Annals of Emergency Medicine, 61(4), 484. https://doi.org/10.1016/j.annemergmed.2013.02.009

National Research Council (U.S.) (Ed.). (2001). Strategic highway research: Saving lives, reducing congestion, improving quality of life. Transportation Research Board, National Research Council : National Academy Press.

Olsson, N. (2023). A validation study comparing performance in a low-fidelity train-driving simulator with actual train driving performance. Transportation Research Part F: Traffic Psychology and Behaviour, 97, 109–122. https://doi.org/10.1016/j.trf.2023.07.007

Papadimitriou, E., Lassarre, S., & Yannis, G. (2016). Introducing human factors in pedestrian crossing behaviour models. Transportation Research Part F: Traffic Psychology and Behaviour, 36, 69–82. https://doi.org/10.1016/j.trf.2015.11.003

Pillai, A. (2017). Virtual Reality based Study to Analyse Pedestrian Attitude towards Autonomous Vehicles. https://aaltodoc.aalto.fi/handle/123456789/28563

Prato, C. G., Bekhor, S., & Pronello, C. (2005). Methodology for Exploratory Analysis of Latent Factors Influencing Drivers' Behavior. Transportation Research Record, 1926(1), 115–125. https://doi.org/10.1177/0361198105192600114

Rao, S., Ghosh, S., Rodriguez, G. P., Röggla, T., Cesar, P., & El Ali, A. (2023). From Video to Hybrid Simulator: Exploring Affective Responses toward Non-Verbal Pedestrian Crossing Actions Using Camera and Physiological Sensors. International Journal of Human–Computer Interaction, 39(16), 3213–3236. https://doi.org/10.1080/10447318.2023.2224955

Rodríguez Palmeiro, A., van der Kint, S., Vissers, L., Farah, H., de Winter, J. C. F., & Hagenzieker, M. (2018). Interaction between pedestrians and automated vehicles: A Wizard of Oz experiment. Transportation Research Part F: Traffic Psychology and Behaviour, 58, 1005–1020. https://doi.org/10.1016/j.trf.2018.07.020

Samanta, S., Mukherjee, A., Ashour, A. S., Dey, N., Tavares, J. M. R. S., Abdessalem Karâa, W. B., Taiar, R., Azar, A. T., & Hassanien, A. E. (2018). Log Transform Based Optimal Image Enhancement Using Firefly Algorithm for Autonomous Mini Unmanned Aerial Vehicle: An Application of Aerial Photography. International Journal of Image and Graphics, 18(04), 1850019. https://doi.org/10.1142/S0219467818500195

Sandbhor, S., & Chaphalkar, N. B. (2019). Impact of Outlier Detection on Neural Networks Based Property Value Prediction. In S. C. Satapathy, V. Bhateja, R. Somanah, X.-S. Yang, & R. Senkerik (Eds.), Information Systems Design and Intelligent Applications (pp. 481–495). Springer. https://doi.org/10.1007/978-981-13-3329-3_45

Sinha, A., Chand, S., Vu, V., Chen, H., & Dixit, V. (2021). Crash and disengagement data of autonomous vehicles on public roads in California. Scientific Data, 8(1), 298. https://doi.org/10.1038/s41597-021-01083-7

Sivak, M. (1998). The information that drivers use: Is it indeed 90 percent visual? The UMTRI Research Review, 29(1), 1.

Stange, V., Kühn, M., & Vollrath, M. (2022a). Manual drivers' experience and driving behavior in repeated interactions with automated Level 3 vehicles in mixed traffic on the highway. Transportation Research Part F: Traffic Psychology and Behaviour, 87, 426–443. https://doi.org/10.1016/j.trf.2022.04.019

Stange, V., Kühn, M., & Vollrath, M. (2022b). Safety at first sight? – Manual drivers' experience and driving behavior at first contact with Level 3 vehicles in mixed traffic on the highway. Transportation Research Part F: Traffic Psychology and Behaviour, 87, 327–346. https://doi.org/10.1016/j.trf.2022.04.004

Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2017). Human Factors Methods: A Practical Guide for Engineering and Design (2nd ed.). CRC Press. https://doi.org/10.1201/9781315587394

Steve, S. (2011). Intelligent Transportation Systems—Integrated Vehicle-Based Safety Systems (IVBSS). https://its.dot.gov/research_archives/ivbss/index.htm

Taelman, J., Vandeput, S., Spaepen, A., & Van Huffel, S. (2009). Influence of Mental Stress on Heart Rate and Heart Rate Variability. In J. Vander Sloten, P. Verdonck, M. Nyssen, & J. Haueisen (Eds.), 4th European Conference of the International Federation for Medical and Biological Engineering (pp. 1366–1369). Springer. https://doi.org/10.1007/978-3-540-89208-3_324

Tan, Z., Dai, N., Su, Y., Zhang, R., Li, Y., Wu, D., & Li, S. (2022). Human–Machine Interaction in Intelligent and Connected Vehicles: A Review of Status Quo, Issues, and Opportunities. IEEE Transactions on Intelligent Transportation Systems, 23(9), 13954–13975. https://doi.org/10.1109/TITS.2021.3127217

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. International Journal of Medical Education, 2, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Thompson, W. R. (1947). USE OF MOVING AVERAGES AND INTERPOLATION TO ESTIMATE MEDIAN-EFFECTIVE DOSE: I. Fundamental Formulas, Estimation of Error, and Relation to Other Methods. Bacteriological Reviews, 11(2), 115–145. https://doi.org/10.1128/br.11.2.115-145.1947

Transport research center. (2024). AV Crashes | Autonomous Vehicle Crashes. Autonomous Vehicle Crashes. https://www.avcrashes.net/

U.S. Department of Transportation Federal Highway Administration. (n.d.). Next Generation Simulation (NGSIM) Open Data. Next Generation Simulation (NGSIM) Open Data. Retrieved June 15, 2024, from https://datahub.transportation.gov/stories/s/Next-Generation-Simulation-NGSIM-Open-Data/i5zb-xe34/

van Beers, W. C. M., & Kleijnen, J. P. C. (2004). Kriging interpolation in simulation: A survey. Proceedings of the 2004 Winter Simulation Conference, 2004., 1, 121. https://doi.org/10.1109/WSC.2004.1371308

Wen, X., Jian, S., & He, D. (2023). Modeling the Effects of Autonomous Vehicles on Human Driver Car-Following Behaviors Using Inverse Reinforcement Learning. IEEE Transactions on Intelligent Transportation Systems, 24(12), 13903–13915. https://doi.org/10.1109/TITS.2023.3298150

Williamson, A., Grzebieta, R., Eusebio, J. E., Zheng, W. Y., Wall, J., Charlton, J., Lenne, M., Haley, J., Barnes, B., Rakotonirainy, A., Woolley, J., Senserrick, T., Young, K., Haworth, N., Regan, M., Cockfield, S., Healy, D., Cavallo, A., Di Stefano, M., … Cornish, M. (2015). The Australian Naturalistic Driving Study: From beginnings to launch. In I. Cameron, N. Haworth, & L. McIntosh (Eds.), Proceedings of the 2015 Australasian Road Safety Conference (ARSC2015) (pp. 1–7). Australasian College of Road Safety (ACRS). http://acrs.org.au/files/papers/arsc/2015/WilliamsonA%20277%20The%20Australian%20naturalistic%20driver%20study.pdf

Wu, M., Wang, N., & Yuen, K. F. (2023). Can autonomy level and anthropomorphic characteristics affect public acceptance and trust towards shared autonomous vehicles? Technological Forecasting and Social Change, 189, 122384. https://doi.org/10.1016/j.techfore.2023.122384

Xu, S. S., Mak, M.-W., & Cheung, C.-C. (2019). Towards End-to-End ECG Classification With Raw Signal Extraction and Deep Neural Networks. IEEE Journal of Biomedical and Health Informatics, 23(4), 1574–1584. https://doi.org/10.1109/JBHI.2018.2871510

Yan, Z., Chakraborty, D., Parent, C., Spaccapietra, S., & Aberer, K. (2011). SeMiTri: A framework for semantic annotation of heterogeneous trajectories. Proceedings of the 14th International Conference on Extending Database Technology, 259–270. https://doi.org/10.1145/1951365.1951398

Yang, J., Rahardja, S., & Fränti, P. (2019). Outlier detection: How to threshold outlier scores? Proceedings of the International Conference on Artificial Intelligence, Information Processing and Cloud Computing, 1–6. https://doi.org/10.1145/3371425.3371427

Zhang, J., Chen, Z., Liu, W., Ding, P., & Wu, Q. (2021). A Field Study of Work Type Influence on Air Traffic Controllers' Fatigue Based on Data-Driven PERCLOS Detection. International Journal of Environmental Research and Public Health, 18(22), Article 22. https://doi.org/10.3390/ijerph182211937

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

Zhao, X., Wang, Z., Xu, Z., Wang, Y., Li, X., & Qu, X. (2020). Field experiments on longitudinal characteristics of human driver behavior following an autonomous vehicle. Tran sportation Research Part C: Emerging Technologies, 114, 205–224. https://doi.org/10.1016/j.trc.2020.02.018

Zhou, A., Liu, Y., Tenenboim, E., Agrawal, S., & Peeta, S. (2023). Car-Following Behavior of Human-Driven Vehicles in Mixed-Flow Traffic: A Driving Simulator Study. IEEE Transactions on Intelligent Vehicles, 8(4), 2661–2673. https://doi.org/10.1109/TIV.2023.3257962

Zhou, M., Qu, X., & Li, X. (2017). A recurrent neural network based microscopic car following model to predict traffic oscillation. Transportation Research Part C: Emerging Technologies, 84, 245–264. https://doi.org/10.1016/j.trc.2017.08.027