# Can Safe Driving Patterns Be Identified? An Exploratory Analysis

Dimitrios I. Tselentis[1], Eleonora Papadimitriou[1](✉), and Arturo Tejada[2,3]

[1] Faculty of Technology, Policy, and Management, Safety and Security Science Section, Delft University of Technology, Delft, The Netherlands
E.Papadimitriou@tudelft.nl
[2] TNO, Mobility and Built Environment Unit, Integrated Vehicle Safety, Helmond, The Netherlands
[3] Dynamics and Control Group, Faculty of Mechanical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands

**Abstract.** In order to improve road safety, recent studies suggest that it is important to study and identify the optimal driving benchmarks that reflect the safest driving behaviour that may be observed by human drivers. The objective of this paper is to identify boundaries of risky and typical driving by studying the car-following driving behaviour. The data used in this study was collected by TNO in a recent naturalistic driving study. The distributions of driving metrics related to the following and leading vehicle were illustrated to understand their shapes and outliers. The safety-related car-following driving metrics of Time to Collision (TTC), Deceleration Rate to Avoid the Crash (DRAC), Crash Index (CI) and over-speeding were calculated, with risky thresholds obtained from the literature, and typical driving thresholds based expert assessors' ratings. Principal Component Analysis (PCA) was applied to these metrics and showed that 'optimal driving' can be represented by one linear component that represents over 95% of the total dataset's variance.

**Keywords:** Car-following · driving metrics · Principal Component Analysis

## 1 Introduction and Literature Review

Driving is a complex task that necessitates a blend of motor and cognitive skills. Drivers must adeptly manage their vehicles in traffic while sustaining critical cognitive functions such as attention, visuospatial coordination, and executive functions (1). Amid this dynamic landscape, drivers engage with their environment and other road users, applying their skills and experience (2). Although most interactions occur successfully, there are sporadic instances of traffic conflicts, near-miss situations, or actual collisions. Over recent decades, extensive research has been conducted on preventing and mitigating these critical moments of increased traffic risk (3). Notably, in recent years, surrogate safety measures have emerged as a valuable tool, enabling proactive assessment of risk before actual accidents occur. This proactive approach to safety management holds potential for averting safety critical events before they happen and permits a proactive strategy based on early risk evaluation (4).


C. McNally et al. (Eds.): TRAconference 2024, LNMOB, pp. 76–82, 2026.
https://doi.org/10.1007/978-3-031-88974-5_12

Nonetheless, in the broader field of safety science, it has been proposed that solely concentrating on negative outcomes offers only a partial comprehension of a system's safety (5). As systems, especially complex ones, continue to evolve, new collision mechanisms may surface, and decomposing the role of individual risk factors becomes more difficult. Conversely, there is valuable knowledge to be gained from closely observing the 'typical' functioning of the system and how users/operators handle inherent risks as part of their everyday activities (6). In order to improve road safety specifically, recent studies suggest that it is important to study and identify the optimal driving benchmarks that reflect the most successful driving behaviour, in terms of safety, that may be observed by human drivers (7). The initial step to achieve this is by defining the boundaries of risky and typical driving in a multivariate framework. This paper aims to identify these boundaries by studying the car-following driving behaviour.

## 2 Data Collection and Methodology

TNO (Dutch Organisation of Applied Research) recently conducted a driving study (the reader is referred to (8) for more details) where the kinematic variables of the ego (test) vehicle (longitudinal and lateral positioning, speed, acceleration, etc.) were recorded, together with the surrounding traffic variables (using a combination of Mobileye, front radars and cameras) of an instrumented vehicle. Fifteen drivers participated in the study, who had more than 7 years of driving experience and drive at least 10.000 km per year. Each driver drove twice in the same location, the stretch of the A2 highway between Best and Boxtel in the Netherlands between exits 25 and 27. Each driver drove for approximately 40 min, given the speed limit (100kph).

One expert driving assessor from the Centraal Bureau Rijvaardigheidsbewijzen (the Dutch agency in charge of awarding driver licenses) accompanied and assessed in real time each driver. The assessors (two in total), were requested to annotate whether driving was "competent", "neutral" or "not competent" by pressing one of three buttons at least once every 30 s. The kinematic data, together with the button presses, were recorded at a 20Hz frequency by the same storage device, the StreetLive box, in ROS format, which was then post-processed into Matlab files. These files were processed in Python for modeling and visualization purposes.

In order to understand the shapes of the driving metrics' distributions and identify the outliers, those were visualized. These distributions included metrics of both the following and leading vehicle as well as their interaction. Metrics encompassed factors like speed, acceleration, Time to Collision (TTC), distance from the leading vehicle, lane positioning, and relative velocity. Subsequently, after excluding lateral movements within a 30-s window around lane changes, safety-oriented car-following metrics were computed. These included TTC, Deceleration Rate to Avoid the Crash (DRAC), Crash Index (CI), and instances of over-speeding. Thresholds denoting risk levels for each of these metrics were determined through prior literature examination and the time spent within the risky area was calculated for each metric.

Having excluded the risky areas from the data, the 'typical driving' area was defined by further filtering the data to identify the proportions of non-risky driving that corresponded to the "competent" button presses by the expert assessor for each metric. In order to reduce the dimension of "typical" driving to one principal component, Principal Component Analysis (PCA) was applied, which is a statistical technique employed to reduce the dimensionality of data while preserving its essential information. It identifies orthogonal axes (principal components) that capture the maximum variance within the dataset. PCA aids in simplifying complex data, enhancing interpretability, and revealing patterns. It is extensively utilized in various fields, including transport safety for the purposes of data compression, feature selection, and visualization, enabling researchers to discern underlying structures and reduce data complexity for further analysis (9).

## 3 Results

### 3.1 Descriptive Analysis of Safety Metrics

Figure 1 shows the distributions of the ego vehicle's speed, acceleration, and lateral position, i.e. the distance of the vehicle's central axis from the lane's right border. The average speed is found to be around 95 kph. It is also observed that slightly more people drive between 95–100 than 90–95 kph. A longer tail can be noticed on the left side of the distribution, which mainly captures driving behaviour while entering or exiting the highway. Regarding the acceleration, which is converted to kph per second, it averages around 0 kph/s. Finally, the positive values present slightly higher frequency than the negative values. With regards to the lane positioning shown in the right panel, the distribution is observed to be symmetric and centered around 1.60 m.

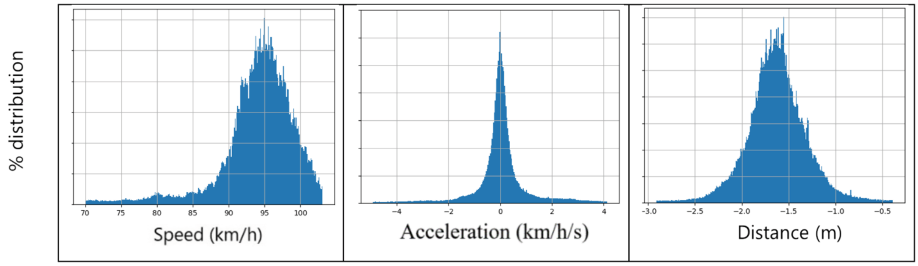


**Fig. 1.** Distribution of speed (left panel), acceleration (middle panel) and lane positioning (right panel) of the ego vehicle of the experiment

Figure 2 shows the distributions of the relative speed between the ego and the leading vehicle, the TTC between the two vehicles, as well as the distance from the leading vehicle. The fact that the right tail of the relative speed distribution is longer and has higher frequency than the left tail, shows a tendency of the following vehicles to approach the leading vehicle. Additionally, the distribution of TTC in the middle panel is negatively skewed with mean value around 6 s. It is observed that a share of critical interactions are recorded with TTC less than 2 s, which according to literature corresponds to critical conflicts. At the same time, a share of interactions with TTC less than 4 s may be considered to reflect the "successful" interactions, while the right part of the distribution may be considered to reflect the "undisturbed" interactions (7). Regarding distance from the vehicle ahead, it can be observed that the distribution is lognormal and positively skewed.

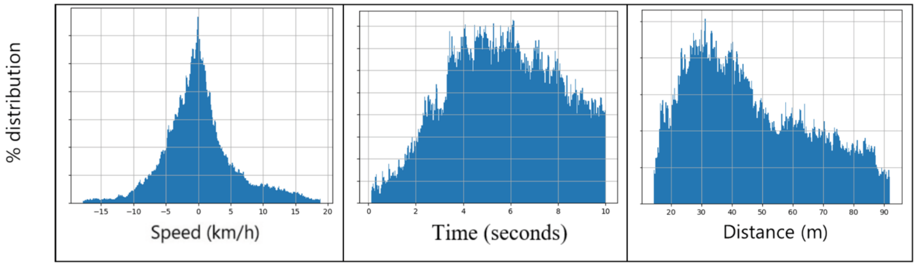


**Fig. 2.** Distribution of relative speed (left panel), TTC (middle panel) and distance to the leading vehicle (right panel) of the ego vehicle of the experiment

Figure 3 shows the breakdown of the relative speed and lane positioning distributions with respect to the button presses of the assessors. It can be seen that the vast majority of small relative speeds between the leading and the ego vehicles are rated as "competent" or "neutral", whereas a small share of large speed difference is rated as "competent", possibly because it would be acceptable due to contextual factors. Similarly in the distance distribution, the vast majority of very low distances were given a "non-competent" rating, and the vast majority of high distances were given a "competent" rating; a small share of close interactions were given a "competent" or "neutral" rating, probably reflecting situations where the ego vehicle engaged in these interactions in a safe way. This descriptive analysis shows that examining a single driving metric may not provide a full understanding of driver's performance. Additionally, the assessment by an experienced evaluator can contribute valuable insights into the safety aspects in each case.

Following the above, it was decided to focus on longitudinal driving and exclude lateral movements occurring 30 s before and after lane changes. The car-following metrics were calculated from the experimental data and their risky thresholds were adopted from the literature (10), as follows: 1.6 s for TTC, 16.95 for CI, 3.35 for DRAC and 100kph for Speed. It is noted that the CI threshold was specified on the basis of the boundary conditions of the formula provided in (10) on the basis of the risky thresholds of the included metrics, i.e. by considering 100 kph speed limit for both the leading and the following vehicle, 0.6 m/s$^2$ acceleration for the following vehicle and 1.6 s TTC for the following vehicle.

### 3.2 Principal Component Analysis for Typical Driving

A PCA model estimated 4 components (based on the 4 metrics). The 1$^{st}$ component represents 99.6% of the total variance of the dataset, meaning that the data can be adequately represented without using the rest of the components. In order to define the 'typical' driving period, the values lying within the 10% and 90% percentiles of this component were retained.

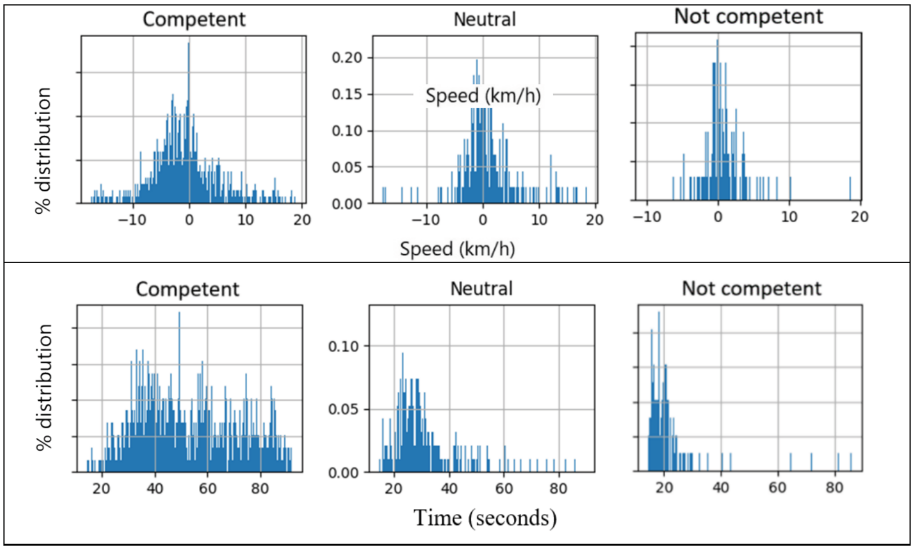


**Fig. 3.** Breakdown of the distribution of relative speed (upper panel), and TTC (lower panel) of the ego vehicle per assessor ratings (competent, neutral, not competent)

## 4 Discussion and Conclusions

The 4 driving metrics used were found to be relevant for distinguishing both safe and typical driving segments. The methodology involved creating a latent construct to define typical driving and applying three criteria: staying within non-risky thresholds, having a 'competent driving' button pressing by an expert assessor, and excluding extreme percentiles. This resulted in a concise representation of typical driving, which may potentially serve as an optimization input. PCA revealed that these driving metrics can be formed into a single linear component of typical driving.

This research has some limitations. The analysis was limited to car-following conditions, and lateral movements were excluded, due to difficulty in estimating appropriate metrics of lateral movement (e.g. gaps, lateral acceleration) from the available data.

The determination of typical driving behavior was founded on a blend of data-driven (non-risky driving time) and expert-based (assessor's "competent" rating) criteria. We acknowledge that neither of these criteria is flawless individually, nor is their combination. Obviously, there may be some subjectiveness in assessors' ratings. On the other hand, the fact that the proposed methodology lies on the existence of expert-based ratings, it is not directly transferable to other existing naturalistic driving datasets.

These results may be useful to researchers on road safety and human factors, as they offer insights into the skills used by drivers to navigate real-world driving and manage risks. Moreover, these findings have may practical implications for developing better Advanced Driver Assistance Systems (ADAS), by learning from the whole spectrum of driver behaviour and developing automation and support strategies that are recognizable and trustworthy by human drivers. Future research may study optimal driving as an optimisation process, in which the "typical" driving performance is maximised, and any risky driving situations is minimised.

**Acknowledgements.** This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101027222.

## References

1. Pavlou, D., et al.: Comparative assessment of the behaviour of drivers with mild cognitive impairment or Alzheimer's Disease in different road and traffic conditions. Transport. Res. F: Traffic Psychol. Behav. **47**, 122–131 (2015)
2. Rasmussen, J.: The role of error in organizing behaviour. Ergonomics **33**(10–11), 1185–1199 (1990). https://doi.org/10.1080/00140139008925325
3. Arun, A., Haque, M.M., Bhaskar, A., Washington, S., Sayed, T.: A systematic mapping review of surrogate safety assessment using traffic conflict techniques. Accid. Anal. Prev. **153**, 106016 (2021)
4. Wang, C., Xie, Y., Huang, H., Liu, P.: A review of surrogate safety measures and their applications in connected and automated vehicles safety modeling. Accid. Anal. Prev. **157**, 106157 (2021). https://doi.org/10.1016/j.aap.2021.106157
5. Hollnagel, E.: Safety-I and Safety-II: The Past and Future of Safety Management, pp. 1–187. ISBN 978–147242306–1, 978–147242305–4, Ashgate Publishing Ltd. (2014)
6. Hollnagel, E., Wears, R.L., Braithwaite, J.: From Safety-I to Safety-II: A White Paper. The Resilient Health Care Net: Published simultaneously by the University of Southern Denmark, University of Florida, USA, and Macquarie University, Australia (2015)
7. Papadimitriou, E., Pooyan Afghari, A., Tselentis, D., van Gelder, P.: Road-safety-II: opportunities and barriers for an enhanced road safety vision. Accid. Anal. Prev., 106723 (2022)
8. Tejada, A., Hogema, J.H., van Dam, E., Souman, J.L., Silvas, E.: StreetProof: monitored deployment for safe and social automated driving. In: Proceeding of the JSAE Annual Congress (Spring), paper s231413, Yokohama, Japan (2023)
9. Tselentis, D.I., Papadimitriou, E., van Gelder, P.: The usefulness of artificial intelligence for safety assessment of different transport modes. Accid. Anal. Prev. Anal. Prev. **186**, 107034 (2023)
10. Tejada, A., Manders, J., Snijders, R., Paardekooper, J.-P., de Hair-Buijssen, S.: Towards a characterization of safe driving behavior for automated vehicles based on models of "Typical" human driving behavior. In: 2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), pp. 1–6 (2020)