Using mobile devices for driving test assessment
a study of acceleration and GPS data

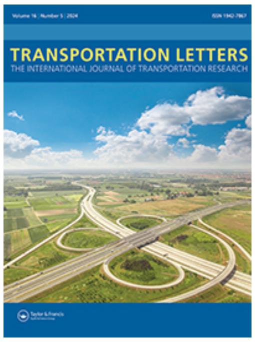
Driessen, Tom; Stefan, David; Heikoop, Daniël; Dodou, Dimitra; de Winter, Joost

DOI
10.1080/19427867.2024.2352198

Publication date
2024

Document Version
Final published version

Published in
Transportation Letters



Taylor & Francis
Taylor & Francis Group

Transportation Letters
The International Journal of Transportation Research

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ytrl20

Using mobile devices for driving test assessment: a study of acceleration and GPS data

Tom Driessen, David Stefan, Daniël Heikoop, Dimitra Dodou & Joost de Winter

To cite this article: Tom Driessen, David Stefan, Daniël Heikoop, Dimitra Dodou & Joost de Winter (16 May 2024): Using mobile devices for driving test assessment: a study of acceleration and GPS data, Transportation Letters, DOI: 10.1080/19427867.2024.2352198

To link to this article: https://doi.org/10.1080/19427867.2024.2352198

TRANSPORTATION LETTERS
https://doi.org/10.1080/19427867.2024.2352198

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Using mobile devices for driving test assessment: a study of acceleration and GPS data

Tom Driessen [a*], David Stefan [b*], Daniël Heikoop [b], Dimitra Dodou [a] and Joost de Winter [a]

[a]Faculty of Mechanical Engineering, Delft University of Technology, Delft, Netherlands; [b]Research & Development, Centraal Bureau Rijvaardigheidsbewijzen, Rijswijk, Netherlands

ABSTRACT
There is a need to improve the validity of the driving test as a measure of an individual's ability to drive safely. This paper explores the use of algorithms to analyze acceleration and GPS data from a smartphone and a GoPro to distinguish between different driving styles, as performed by experienced examiners portraying stereotypical driving test candidates. Measures from nine driving tests were analyzed, including (harsh) accelerations, jerk, mean speed, and speeding. Results showed that the type of car, instructed driving style, and driving route impacted the dependent measures. It is concluded that GPS and accelerometer data can effectively distinguish between cautious, normal, and aggressive driving. However, it is important to consider additional sensors, such as cameras, to allow for more context-aware assessments of driving behavior. Furthermore, we demonstrate methods to quantify variations in road conditions and provide suggestions for presenting the data to driving examiners.

ARTICLE HISTORY
Received 9 March 2023
Accepted 2 May 2024

KEYWORDS
Driving test; driver education; sensor measurements; acceleration; driving styles; cautious driving; aggressive driving; driver monitoring

Introduction

Young drivers face a high risk of road accidents due to a combination of factors such as inexperience and limited skills as well as immaturity and risk-taking behavior (Lajunen, Sullman, and Gaygısız 2022; Rolison and Moutari 2020; Weast and Monfort 2021). To address this issue, several countermeasures have been implemented. One such measure is graduated driver licensing (GDL), which restricts the driving privileges of new drivers in stages as they gain experience (Curry et al. 2017; Fell et al. 2011; Poirier, Blais, and Faubert 2018; Williams 2017). Another countermeasure is the introduction of new vehicle technology, such as front crash prevention and blind spot monitoring, which can help reduce the accident risk of young drivers in particular (Mueller and Cicchino 2022). Effective enforcement of traffic laws (Bates et al. 2020; De Waard and Rooijers 1994) and anti-speeding and anti-drink-drive education campaigns can further reduce the number of accidents involving young drivers (Tay 2005). Finally, the driving test is considered an important screening mechanism that helps ensure that only drivers who are deemed skilled receive their driver's license.

The driving test is often the only formal evaluation of a person's driving skills before they are granted a driver's license (Helman et al. 2017). However, the driving test may not provide a veridical assessment of a person's driving abilities, as it only provides a snapshot of the candidate's skills. The likelihood of making mistakes during the test can be influenced by external factors, such as weather conditions and the occurrence of specific situations on the road. Furthermore, even though driving examiners are trained and qualified, there is still room for subjectivity and human error or inconsistency in their verdict (Baughan et al. 2005). Another challenge in driver testing is the fact that some candidates are disagreeable or may even become aggressive when they hear that they failed the exam (Alsharef, Albert, and Bhandari 2022; Foxe 2020). Hence, there is a need for a more data-driven presentation of the test verdict.

In a previous study with driving examiners in the Netherlands, it was found that examiners would like to have access to data-based evidence to support their decisions to pass or fail a candidate (Driessen et al. 2021). Examples mentioned by the examiners include dashcam footage, recordings of the candidate's viewing behavior, and data on speed, headway, and braking behavior. The examiners indicated they wanted to be able to access such data in a raw (e.g. graphs, footage) or semi-processed (e.g. good/bad evaluations) form so that they could provide more detailed explanations for their verdict. However, the examiners also believed that current technology is not advanced enough to fully replace human judgment with an automated pass-fail system, indicating that technology can assist but not replace the human evaluator (Driessen et al. 2021).

In the area of usage-based insurance, devices like mobile phones and dongles are widely used to monitor driving behavior. These devices can record driving measures such as speed and acceleration and offer the advantage of not requiring modifications to the vehicle or special hardware installation. Studies have shown that hard braking is a reliable predictor of accident risk for car drivers (Hunter et al. 2021; Ma et al. 2018; Stipancic, Miranda-Moreno, and Saunier 2018) and truck drivers (Cai et al. 2021; Driessen et al. 2024). Additionally, studies have explored the use of mobile phones to identify different driving styles, such as dangerous and aggressive driving (Carlos et al. 2020; Chan et al. 2020; Johnson and Trivedi 2011; Othman et al. 2022). Research has also explored the use of smartphone apps for providing personalized feedback to drivers (Marafie et al. 2021) and stimulating their receptivity for feedback by means of gamification techniques, such as leaderboards, rewards, and group forming (Musicant and Lotan 2016; Shanly et al. 2018). Nambi et al. (2019) demonstrated several techniques for measuring maneuvers during Indian driving tests and claimed success using driver gaze monitoring to detect mirror scanning before lane

CONTACT Joost de Winter j.c.f.dewinter@tudelft.nl Faculty of Mechanical Engineering, Delft University of Technology, Delft, Netherlands
*Joint first authors.

changes. They further used a combination of camera, inertial, and GPS data for trajectory tracking.

Despite the widespread use of sensor measurements for driver assessment, there is limited research examining the validity of these methods from an algorithmic viewpoint and in such a way that it can be applied to the on-road driving test. Our study aims to fill this gap by presenting a series of algorithms for evaluating driving performance in these tests. These driving tests were carried out by experienced driving examiners, who emulated typical driving styles encountered during exams. The algorithms are explained in a step-by-step manner, allowing others to use them, and the code for this work is provided as supplementary material.

Methods

The data were collected in cooperation with the driving examiner training center of the Dutch Central Office of Driving Certification (CBR) in Leusden, Netherlands. At this training facility, driving examiner trainees are trained to become licensed driving examiners. A part of their training consists of on-road sessions, in which qualified instructors (active or former driving examiners) emulate driver behaviors commonly encountered at the driving test. The examiner trainee in the passenger seat receives no information about the role the driver takes and is expected to take the role of a real driving examiner and form a pass or fail verdict based on the acted driving style.

Data collection

The observations took place during 21 training sessions between March 30, 2022 and April 13, 2022. All drivers involved were asked for consent before the start of the experiment. The research was approved by the Human Research Ethics Committee of the Delft University of Technology, approval number 2302.

Acceleration data were recorded using the smartphone app Matlab Mobile version 9.1.2 (Mathworks 2021) at a frequency of 10 Hz on an iPhone X (model A1865) and stored on the smartphone's local drive. The phone was placed on the backseat, with its back part fixed between the backrest and the seating area of the seat, and with its longitudinal axis and the car's longitudinal axis aligned. The screen faced upward, and the charging port pointed to the back of the car. Additionally, a GoPro Max was used to record video of the road ahead (1920 × 1080 pixels at 30 Hz). The video files contained embedded accelerometer recordings (at about 200 Hz) and GPS data (at about 17 Hz). These data logs were extracted from the video files using goprotelemetryextractor.com (Telemetry Overlay S.L. 2022). The Appendix shows several example rows of data for both devices.

Driving tests

The drives all emulated a standard driving exam conducted by the CBR, having a duration of approximately 30 minutes. The drives-started and ended at the same CBR location and sometimes involved driving on the same road segments. However, the drivers (i.e. 'test candidates') drove different routes, as the routes in Dutch driving tests are not set in advance but rather are determined by the examiner (in our study: the examiner trainee), based on factors such as traffic conditions and road closures.

The 21 driving tests emulated various driving styles, including 'a good driving candidate,' 'a candidate who was close to passing or failing due to certain mistakes,' 'a good candidate with poor viewing behavior or timing of actions,' 'a slow candidate,' 'a nonchalant candidate,' 'a fast candidate,' etc. Before each drive, the driver received a sheet containing the role description for the current ride and the intended result (pass/fail). The examiner in the passenger seat was blind to the instructed driving style.

From the 21 driving tests, we selected a total of nine driving tests (3 per car) because they allowed for systematic comparison. The other driving tests were of limited validity for further analysis because of various issues (e.g. inconsistent phone placement between drives in the same car, or an interruption of the drive). An overview of the nine selected driving tests is provided in Table 1. Each car was driven by a different driver, so there were a total of three drivers.

Data processing

The first step in processing data was to rotate the accelerometer data. Though care was taken to ensure that the GoPro and the phone's x- and y-axes were aligned with the car's frame, the devices still had non-negligible pitch and roll angles relative to the earth, which had to be corrected for.

We computed the orientation of the device (phone or GoPro) from the acceleration measurements in the three perpendicular directions (Figure 1). First, we computed the mean acceleration values in each direction over the entire drive. Next, the orientation of the device was computed using equations for determining orientation relative to the earth's gravitational field (Pedley 2013). Our assumption here is that, although the car is moving and hence continuously experiencing accelerations, the accelerations due to vehicle motion can be expected to average out across the entire drive, leaving just the acceleration component caused by gravity. Next, we computed the roll and pitch angles of the device using arctangent functions (see Pedley 2013, Eqs. 25 & 26). Then, a rotation matrix was computed, which was used to rotate the original acceleration measurements to their new orientation, aligned with the earth's downward gravitational field.

It is noted that the yaw angle is undetermined since it cannot be inferred based on the gravitation vector. In our calculations, the yaw angle with respect to the car was assumed to be 0 deg, which is

Table 1. Nine driving tests used in the analysis.

No	Car	Weather	Date & Time	Emulated driving style of 'candidate'	Emulated test result
1	Peugeot 308 SW 2014	Rainy	4 Apr, 10:13	Difficulty with position on road	Fail
2	Peugeot 308 SW 2014	Rainy	4 Apr, 13:11	Difficulty with vehicle control and steady steering	Fail
3	Peugeot 308 SW 2014	Rainy	4 Apr, 13:58	Inappropriate timing; acting too early/late	Fail
4	Volkswagen T-Roc 2015	Sunny	11 Apr, 10:50	Inappropriate viewing behavior, engine stalling, position on road	Fail
5	Volkswagen T-Roc 2015	Sunny	11 Apr, 13:06	Aggressive/dangerous driving	Fail
6	Volkswagen T-Roc 2015	Sunny	11 Apr, 14:01	Desirable driving, but one large error (merging without looking)	Pass
7	Seat Ateca 2016	Cloudy	13 Apr, 09:21	Cautious/slow driving	Fail
8	Seat Ateca 2016	Cloudy	13 Apr, 10:06	Negligent viewing behavior	Fail
9	Seat Ateca 2016	Cloudy	13 Apr, 11:00	Desirable driving style, but occasional inappropriate looking	Pass

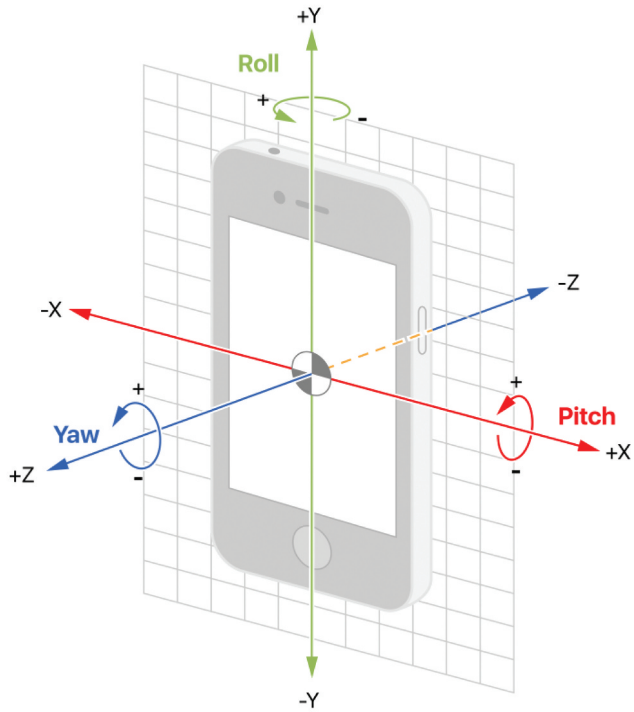


Figure 1. Phone coordinate system (Apple Inc 2022).

a valid assumption since the GoPro and phone were positioned in this manner. Note that selecting another yaw angle will not change the accelerations in the rotated vertical direction (z), but will affect how the accelerations are distributed along the rotated x and y directions (while not changing their combined magnitude).

The measurement of acceleration in a moving vehicle is complicated by high-frequency vibrations caused by uneven roads and engine vibrations. How this noise protrudes in the signal depends on device placement (e.g. hard or soft surface) and the vehicle's damping properties.

A second-order Butterworth zero-phase filter with a cutoff frequency of 0.5 Hz was used to remove these vibrations, resulting in a smoother and more accurate representation of the car's acceleration (Figure 2). Figure 2 illustrates the effect of the filter on the rotated acceleration data in the y-direction, which represents the longitudinal direction of the car. The GPS speed of the GoPro is also shown in the figure.

Speed limit extraction

Speed limits were obtained using the Map Matching API provided by Mapbox (2023). This service takes a driven GPS path and returns the coordinates of the route that was most likely driven, including the speed limits on these roads. To obtain a robust response, the data were first downsampled to a sample rate of 5 s between points, as advised in the API documentation. The paths were split into sections of 100 points each (the maximum number of coordinates allowed per request) and then merged after receiving the speed limits from the API. Then, the coordinates that were left out due to the resampling received the same speed limit of the nearest neighbor from the downsampled set. Upon visual inspection, it was found that the speed zones were correctly assigned, including short exceptions in residential districts, such as school areas. The API requests and processing were done using a Python script that is provided in the supplementary material.

Measures

After the above data pre-processing, five measures were calculated for each of the nine driving tests:

(1) Macc: Mean absolute acceleration in the combined x and y directions (m/s^2). The rotated and filtered longitudinal (y) and lateral (x) accelerations were combined using the Pythagorean theorem.
(2) Mjerk: Mean absolute jerk in the combined x and y directions (m/s^3). The rotated and filtered longitudinal (y) and lateral (x) accelerations were combined using the Pythagorean theorem. Next, the derivative was computed (i.e. jerk in m/s^3), and the mean absolute value was taken. The jerk can be seen as a measure of the abruptness of changing acceleration, and has been previously used in driver assessment (De Groot, De Winter, and Mulder

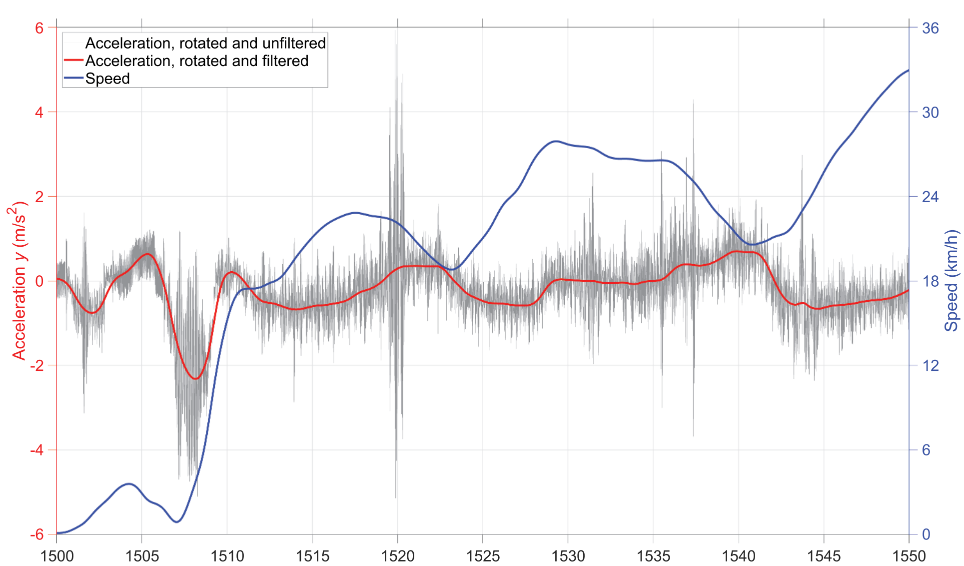


Figure 2. Illustration of the effect of low-pass filtering of the accelerometer data of the GoPro for a portion of Driving Test 1. The figure also shows the low-pass filtered vehicle speed recorded using the GPS of the GoPro. A negative acceleration means that the car is accelerating.

2011; Feng et al. 2017; Itkonen et al. 2017). It has been found to be associated with tailgating and traffic violations, self-reported accident involvement (Bagdadi and Várhelyi 2011), and recorded culpable crashes (Khorram, Af Wåhlberg, and Tavakoli Kashani 2020). Figure 3 illustrates the meaning of jerk, where between 1505 and 1510 s, the driver accelerates; the onset and offset of the acceleration are accompanied by peaks in jerk.

(3) Mspeed: Mean speed (km/h)

(4) MspeedE: Speed limit exceedance (proportion of driving time)

(5) HarshA: Harsh acceleration events, defined as the mean number of combined x and y acceleration threshold exceedances per hour of driving (# per hour). This measure was obtained by identifying all peaks in the combined acceleration and counting the number of peaks that exceeded a threshold value of $3\,\mathrm{m/s^2}$. In the literature, there is no consensus about which threshold to choose (e.g. Khorram, Af Wåhlberg, and Tavakoli Kashani 2020; Stipancic, Miranda-Moreno, and Saunier 2016). Depending on the application and sample size, different threshold values may need to be adopted. Selecting a low threshold will yield a large number of threshold exceedances, which may reflect driving style but may also involve false positives, such as accelerations due to road unevenness. Selecting a high threshold, on the other hand, risks missing important events and will reduce statistical power. After inspection of the acceleration signal, we opted for a threshold of $3\,\mathrm{m/s^2}$. Indicatively, longitudinal decelerations of up to $3\,\mathrm{m/s^2}$ are perceived as 'reasonably comfortable' (Harwood 1992, p. 41).

The accelerometer-based measures (Macc, Mjerk, HarshA) were computed for both the phone and GoPro, while the GPS-derived measures (Mspeed, MspeedE) were computed only for the GoPro. The reason for relying on the GoPro's GPS measurement was that it was more accurate. During several drives, the phone's receiver lost connectivity to the GPS satellites.

Data samples with a GPS GoPro speed below 3 km/h, indicative of the car being stationary or near-stationary, were excluded from the above driver assessment score computations. This was done as such instances, which may include special maneuvers or waiting at a traffic light, do not provide a valid representation of driving abilities.

Results

Table 2 displays the nine driving tests and the corresponding dependent measures. Firstly, it seems the type of car used in the test has an impact on the results. Specifically, Driving Tests 4 to 6, conducted in a Volkswagen, are distinct from the tests performed in a Peugeot (Driving Tests 1 to 3) or a Seat (Driving Tests 7 to 9), where the results of Driving Tests 4 to 6 show relatively high values for the mean absolute acceleration (Macc), mean absolute jerk (Mjerk), and harsh acceleration rate (HarshA).

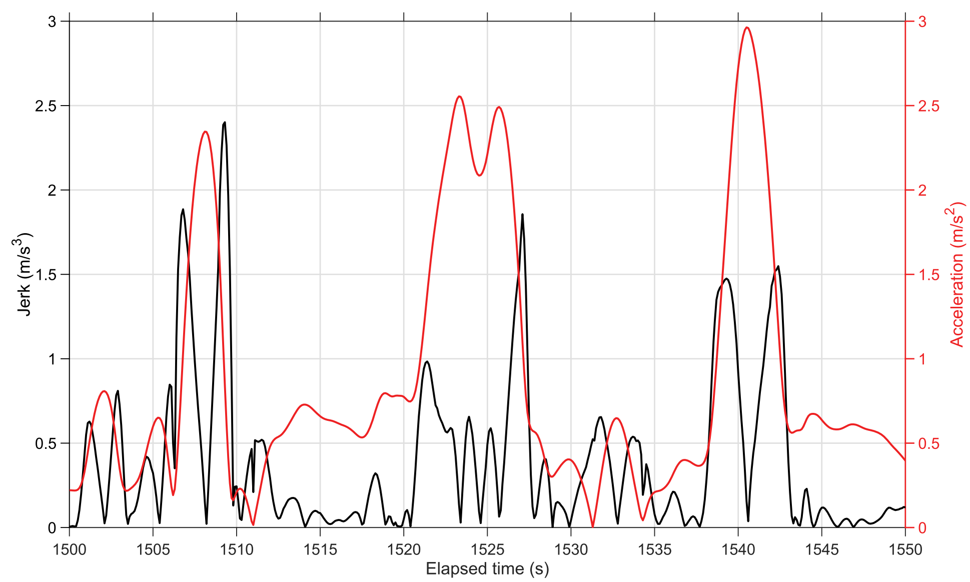


Figure 3. Jerk based on combined acceleration in the *xy*-plane for a portion of Driving Test 1. The selected time interval is the same as shown in Figure 2.

Table 2. Dependent measure scores for the nine driving tests.

		Phone			GoPro				
No	Emulated driving style of 'candidate'	Macc	Mjerk	HarshA	Macc	Mjerk	Mspeed	MspeedE	HarshA
1	Difficulty with position on road	0.86	0.35	35.4	0.81	0.34	36.12	0.08	22.9
2	Difficulty with vehicle control and steady steering	0.78	0.32	34.7	0.73	0.32	36.89	0.10	23.1
3	Inappropriate timing; acting too early/late	0.77	0.33	30.1	0.69	0.31	36.12	0.13	23.2
4	Inappropriate looking, engine stalling, position on road	0.82	0.38	43.2	0.81	0.38	37.31	0.12	40.4
5	Aggressive/dangerous driving	0.81	0.41	54.4	0.81	0.41	47.72	0.43	53.9
6	Desirable driving, but one large error (merging without looking)	0.86	0.37	43.8	0.87	0.37	38.32	0.09	48.1
7	Cautious/slow driving	0.65	0.22	10.5	0.59	0.21	32.61	0.04	8.4
8	Negligent viewing behavior	0.77	0.27	33.3	0.72	0.26	37.03	0.11	16.7
9	Desirable driving style, but occasional inappropriate looking	0.82	0.30	41.0	0.73	0.29	38.35	0.08	28.9

Note. Color coding is applied per column from blue (lowest value) to white (median) to red (highest value).

Secondly, the instructed driving style seems to have an impact on the driving measures. Specifically, Driving Test 7, which was performed with a cautious driving style, was characterized by low scores on all measures compared to Driving Tests 8 and 9, which were conducted in the same car. Driving Test 7 had minimal speeding and was characterized by very few harsh acceleration events. Moreover, Driving Test 5, which was performed with an instructed aggressive/dangerous driving style, had high scores on most of the measures compared to Driving Tests 4 and 6. An exception was the mean absolute acceleration (Macc), which was relatively low, at 0.81 m/s^2 for both the phone and GoPro.

We suspect that the driving route also impacted the measures observed. This is illustrated in Figure 4, which presents the absolute jerk in the *xy*-plane during a portion of Driving Test 5. While all the other driving tests were conducted in environments consisting mostly of roads with speed limits of 30 km/h, 50 km/h, and 100 km/h highways, the driver in Driving Test 5 chose a route through rural areas, primarily consisting of 60 km/h roads. Although the driver was tasked to drive with an aggressive driving style, there were often limited opportunities for aggressive driving other than exceeding the speed limit. That is, the driver drove most of the time on 60 km/h roads consisting of smooth asphalt, sometimes following behind another vehicle (between 1100 and 1400 s). There were some exceptions, such as the 800–1100 s interval, where the driver entered a small village. During these moments, the instructed driving style of the driver became more manifest, as shown by the spikes in the jerk.

Three measures, namely mean combined acceleration, mean absolute jerk, and mean number of harsh accelerations per hour of driving, were calculated using both the accelerometers in the phone and the accelerometers in the GoPro (Table 2). These measures were computed for the nine driving tests, and the results showed a high correlation between the two devices for all three measures ($r = 0.929$, $r = 0.996$, and $r = 0.891$, respectively). The correlations for mean combined acceleration and mean absolute jerk are illustrated in Figure 5.

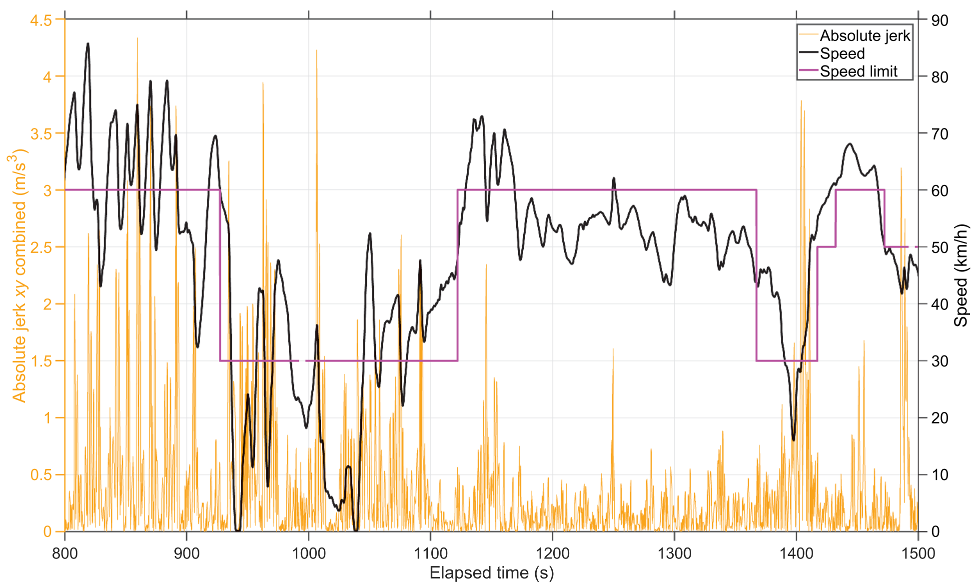


Figure 4. Mean absolute jerk in the *xy*-plane, vehicle speed recorded using GPS, and the speed limit for a portion of Driving Test 5.

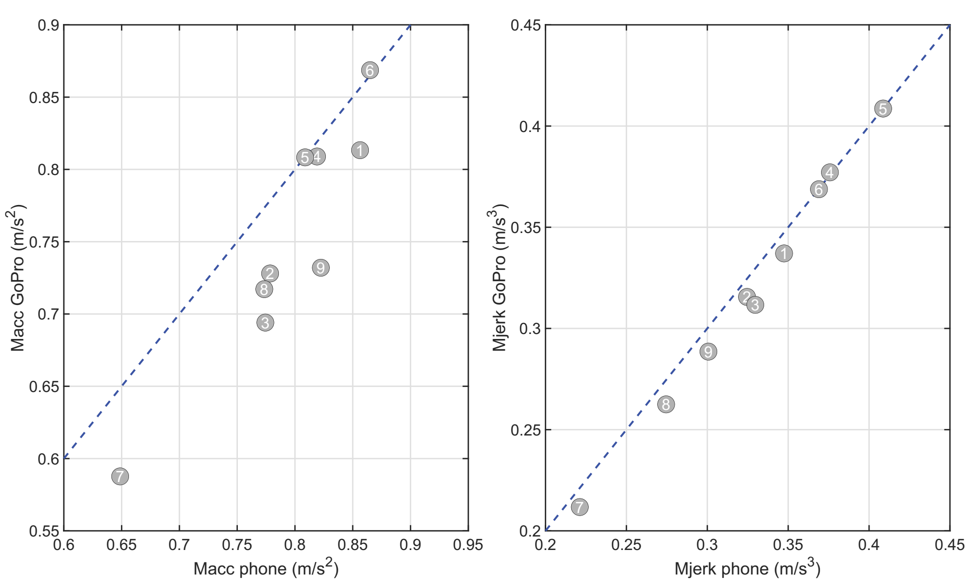


Figure 5. Scatter plot of mean absolute acceleration in the *xy*-plane (left) and mean absolute jerk in the *xy*-plane (right) for the GoPro versus the phone. The dashed line represents the line of unity. Each marker represents a driving test number.

Characterizing the route driven

As evidenced above, the evaluation of driving proficiency using accelerometers is influenced not only by the driving style of the candidate but also by the opportunities for high acceleration that are contingent upon the type of road being driven. Different road conditions require different driving behaviors: Driving on a straight road is different from driving on a road with multiple curves. This raises the issue of how to account for such variations in road conditions. Here, we draw upon prior research that used instrumented vehicles (such as Melman et al. 2021), which indicates that the assessment of driving behavior should be specific to a location, rather than relying on measures from the entire drive.

The number and curvature of curves can indicate the complexity of the road conditions (and the driver's ability to handle these conditions). The curves were extracted using the GPS measurements. First, the bearing of the car was computed from all subsequent GPS coordinates, assuming the earth is a sphere with a radius of 6371 km. The bearing was computed only if the vehicle speed exceeded 5 km/h (at low speeds, the distance between GPS points became too small to determine the bearing reliably). The bearing angle was filtered with a median filter (time window: 2 s), the gaps in the data caused by GPS speeds below 5 km/h were linearly interpolated, and a low-pass Butterworth filter (cutoff frequency of 0.5 Hz) was applied. The effect of the filtering is shown in Figure 6, which compares the bearing before and after filtering. The bearing was differentiated to obtain bearing rate. To prevent abrupt jumps in the angle due to its limited range between 0 and 360 degrees, the unwrap function was used before differentiating, which replaces jumps greater than 180 degrees by their 360-degree complement, resulting in a continuous line. As differentiating the data amplifies noise, the bearing rate was filtered using a median filter and a Butterworth filter with the same parameters as mentioned above. Finally, the curvature of the car's path was computed by dividing the bearing rate by the momentary GPS speed. To further reduce any noise in the curvature data, a Butterworth filter with a cutoff frequency of 0.5 Hz was applied.

We extracted the peaks in the curvature data to count the number of curved paths of the car as well as the moments of those curves. Then, several route statistics were computed, which were tabulated in Table 3:

- Proportion of time driven under each speed limit (30, 50, 60, 80, or 100 km/h). These five values were normalized so that the total is equal to 100%.

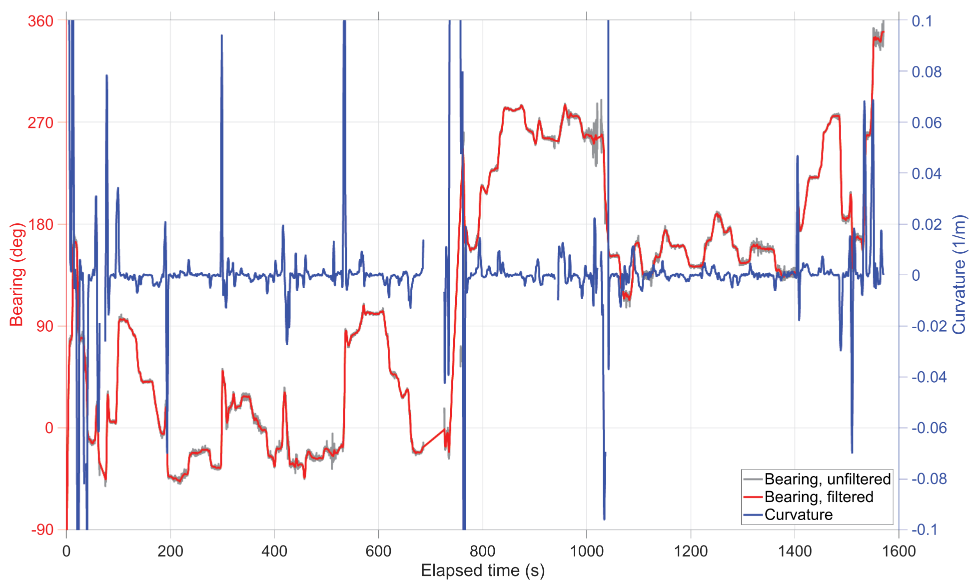


Figure 6. Calculated bearing before and after filtering, and path curvature, for Driving Test 5. A bearing angle of 90 deg corresponds to driving northbound, an angle of 180 corresponds to driving westbound, etc. The *y*-axis for the curvature was constrained to −0.1 and 0.1, corresponding to a turn radius of 10 m. Note that high or low curvature values occurred when the vehicle was driving slowly (see Figure 4).

Table 3. Route statistics, computed from the GoPro GPS data.

No	Emulated driving style of 'candidate'	30 km/h	50 km/h	60 km/h	80 km/h	100 km/h	MildC#	SharpC#	MildC Macc	SharpC Macc
1	Difficulty with position on road	0.28	0.53	0.04	0.06	0.10	114.3	43.7	1.18	2.26
2	Difficulty with vehicle control and steady steering	0.27	0.50	0.05	0.06	0.12	101.6	48.5	0.94	1.99
3	Inappropriate timing; acting too early/late	0.33	0.47	0.05	0.07	0.08	108.8	39.4	0.93	1.90
4	Inappropriate looking, engine stalling, position on road	0.30	0.47	0.04	0.07	0.12	130.8	45.2	1.07	2.87
5	Aggressive/dangerous driving	0.20	0.11	0.69	0.00	0.00	90.6	17.1	1.65	2.68
6	Desirable driving, but one large error (merging without looking)	0.32	0.38	0.04	0.01	0.26	101.6	61.5	1.30	2.35
7	Cautious/slow driving	0.29	0.49	0.05	0.06	0.11	125.9	35.7	0.74	1.81
8	Negligent viewing behavior	0.26	0.42	0.06	0.11	0.15	130.5	50.0	1.05	1.77
9	Desirable driving style, but occasional inappropriate looking	0.30	0.25	0.22	0.06	0.18	118.0	48.2	1.26	2.01

Note. Color coding is applied per column from blue (lowest value) to white (median) to red (highest value).

- MildC#: The number of mild curves (absolute curvature between 0.005 and 0.05) per hour of driving. This corresponds to a turn radius between 20 and 200 m.
- SharpC#: The number of sharp curves (absolute curvature of 0.05 or greater) per hour of driving. This corresponds to a turn radius of 20 m or less, and can be seen as turning at an intersection, turning around, etc.
- MildC Macc: Mean absolute acceleration in the combined x and y directions (m/s^2), averaged across the mild curves.
- SharpC Macc: Mean absolute acceleration in the combined x and y directions (m/s^2), averaged across the sharp curves.

The statistics for Driving Test 5, referred to as the 'aggressive/dangerous' drive, show that 77% of the drive took place in a 60 km/h zone. Although the number of curves was low, the acceleration in these curves was relatively high compared to other driving tests. This highlights the importance of presenting driving examiners with both objective performance measures (as shown in Table 2) and route statistics (as shown in Table 3) in order to provide a more complete understanding of the driver's behavior. The combination of these two tables makes it clear that the driver in Driving Test 5 was driving aggressively in relatively easy road conditions.

Discussion

We presented algorithms that could help distinguish between overcautious, normal, and aggressive driving during the driving test. We solely relied on accelerometer and GPS data and found that these sensors were enough to identify the overcautious and aggressive driving styles. The percentage of driving time exceeding the speed limit, mean jerk, and mean harsh acceleration rate were effective measures in this discrimination. However, mean absolute acceleration across the entire drive was not a clear indicator, as it can vary greatly depending on the eventfulness of a drive, such as the presence of curves. To overcome this issue, we proposed additional measures, namely the speed limit distribution, mild and sharp curve rate, and mean absolute acceleration in curves, to assess the route driven.

The current study provides several insights into the use of accelerometers and GPS. One of our observations was that the combination of x and y acceleration was found to be robust. In particular, the mean absolute jerk measurement demonstrated a particularly high consistency between a smartphone and a GoPro ($r = 0.996$; see Figure 5, right), even though they employed a different measurement unit and were positioned differently in the car (flat on the back seat vs. upright on the dashboard). The robustness of the jerk measure could be attributed to it reflecting changes in acceleration and thus being less susceptible to possible offsets in the acceleration measurement.

Previous interviews revealed that driving examiners could benefit from data-driven support, particularly in communicating their evaluation to test candidates (Driessen et al. 2021). The current study demonstrates that it is possible to generate numerical scores that reflect the driving style and the dynamic nature of the route. These scores could be presented in graph form or in a map format, for example, by visualizing the GPS-recorded route along with color-coded speed, deviation from the speed limit, harsh accelerations, or moving averages of accelerations or jerk values. If dashcam images are used, it may also be feasible to automatically identify and replay segments in the video where driving behavior was particularly noteworthy (e.g. around the moment of the highest recorded acceleration or jerk). Such techniques have the potential to aid the examiner in explaining their evaluation.

The current study highlights the effectiveness of using GPS and accelerometers to distinguish between slow and fast driving styles. Some drivers were instructed to exhibit poor viewing behavior but otherwise normal driving behavior. Indeed, their driving behavior, as measured by acceleration and speed, appeared normal (Table 2: Driving Tests 4, 8, 9). Incorrect viewing behavior is a common cause of failing the driving exam (De Winter et al. 2008; UK Government 2022), and while it is possible that poor anticipation skills may manifest as harsh accelerations and high jerk (Fisher et al. 2002; Parmet et al. 2015), this relationship is only indirect. Other types of sensors may have to be explored to support the assessment of a candidate's viewing behavior. For example, eye-tracking technology is feasible: eye-tracking systems that detect visual distraction are becoming available in modern cars (e.g. DS Automobiles 2023), and several recent research studies have used eye-tracking in combination with object detection to establish at which object the driver was looking (Kim et al. 2020; Qin et al. 2022).

As previously noted, accelerometer and GPS data alone offer a limited perspective on a driver's performance as they fail to capture the driver's interaction with other road users. To gain a better understanding of driving behavior, object detection based on camera images, similar to those employed by automated vehicles, may be necessary (see Figure 7 for an illustration). An online experiment has revealed that the number of identified individuals and the bounding boxes surrounding other road users can predict perceived risk (De Winter et al. 2023). Figure 8 provides examples of how computer vision techniques could be used, namely by counting the number of persons and estimating headway to the car in front using the width of the bounding box (for more, see Rezaei, Azarmi, and Mir 2021). Automated identification of high-risk scenarios, such as passing another road user too closely, may

Figure 7. Bounding boxes generated using the YOLOv4 algorithm for Driving Test 5 (left: elapsed time = 1039 s, right: elapsed time = 1400 s). A YOLOv4 model (Bochkovskiy, Wang, and Liao 2020) pretrained on the COCO dataset was used (Lin et al. 2014; obtained from sbairagy-MW 2021).

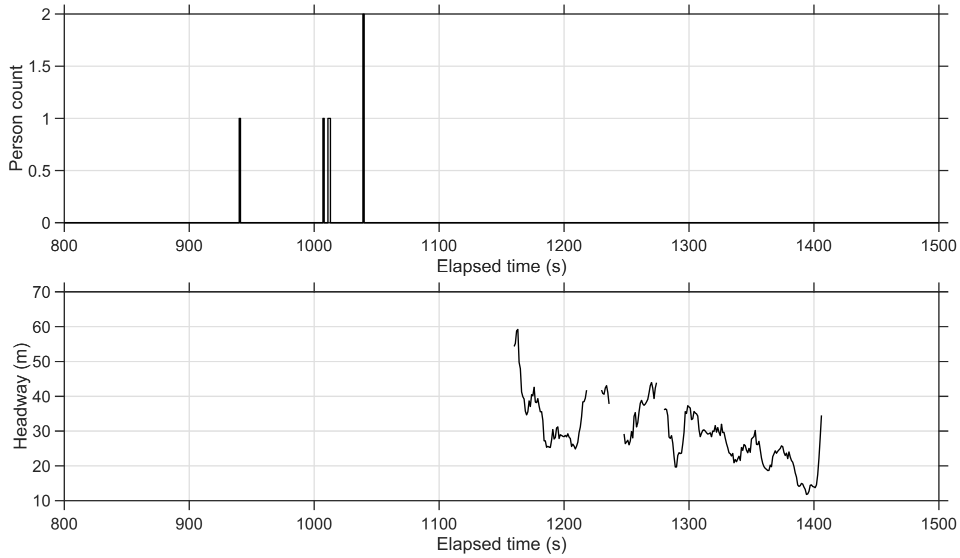


Figure 8. Number of persons (top) and headway to the vehicle in front (bottom) measured at a frequency of 1 Hz for a portion of Driving Test 5 (same portion as shown in Figure 4). In this 700-s interval, the driver encountered a number of persons (see Figure 7, left), proceeded to a 60 km/h road, and followed another car for approximately 250 s (see Figure 7, right). Only objects straight in front of the ego-vehicle were considered (top figure: a 400-pixel horizontal range, bottom figure: a 50-pixel horizontal range. Gaps in the headway of 5 s or less were spline-interpolated.

help examiners form a more objective assessment of driving behavior. However, this type of approach toward the driving test would require further research and validation.

A potential issue in the driver assessment process is that the examiner in the passenger seat occasionally applied the secondary brake pedal to intervene in dangerous situations, which complicates the analysis of the accelerometer data. Future investigations may need to record brake pedal inputs of both the driver (i.e. candidate) and the passenger (i.e. examiner), to isolate their respective contributions. This could also aid in the debriefing session following the driving test, where the examiner could replay the moments of intervention that occurred during the drive, as identified automatically based on the examiner's brake pedal inputs.

A limitation of the current study is its small sample size, comprising nine driving tests and three drivers. In studies analyzing naturalistic driving data, larger sample sizes are typically recommended to ensure that external factors such as weather impact and traffic variability over time are adequately represented across the sample. It is advisable to augment the current analysis with larger sample sizes in future research. Despite the small sample size, this study can serve as an initial framework for designing algorithms aimed at detecting anomalous driving styles typically associated with novice or trainee drivers.

A second limitation is that the study assumed that experienced driving instructors are able to realistically imitate the driving styles of test candidates. Additionally, the driving examiners may have focused on specific, extreme scenarios, which may not be representative of typical driving tests. For example, it has been argued that the driving test is a test of driving skill rather than driving style, and that individuals attempting to obtain their driving license are unlikely to engage in e.g. excessive speeding (Alsharef et al. 2021; Senserrick and Haworth 2005). Another limitation is that traffic density in the test region was relatively low compared to dense city environments.

Apart from driver testing, we see further use in driving data collection during driver training, either for student drivers prior to obtaining their driver's license (Driessen et al. 2021) or for more experienced drivers who engage in self-coaching (Takeda et al. 2012). Similarly, a government report on reforming the Dutch driver education system by Roemer (2021) pointed out the value of using data to help students reflect on their learning progress. Information gathered prior to the final exam may assist driving schools in determining the candidate's readiness for the driving test, thereby reducing the number of unsuccessful test takers (Alsharef, Albert, and Bhandari 2022). Additionally, at present, there is no motivation or requirement for individuals to maintain their driving skills after obtaining a license. The use of data could potentially address this issue.

To determine if recorded data can predict driving test results, collecting data from more driving tests is recommended. For example, sensor data, dash-camera images, and map data could be fed to a machine learning algorithm that predicts pass or fail outcomes. However, it is noted that predicting the test outcome may be challenging due to the fact that test candidates tend to apply for the exam when they have just that amount of driving experience where they have a moderate probability of passing (Baughan et al. 2005).

Conclusions and outlook

The study used accelerometer and GPS data to distinguish between slow, normal, and aggressive driving during driving tests. The findings show that these sensors are sufficient to identify different driving styles, and that the percentage of driving time exceeding the speed limit, mean jerk, and mean harsh acceleration rate are effective measures in this discrimination. However, the study also highlights the limitations of using these sensors alone, as they fail to provide insight into the driver's viewing behavior and interaction with other road users. Future investigations may address this issue by incorporating computer vision methods.

The study concludes that the use of GPS and accelerometers has the potential to aid driving examiners in their assessments and communication with test candidates. However, more research is needed, as the number of driving tests was small and there are limitations associated with experienced driving instructors imitating the driving styles of test candidates. Instead of using ex-examiners, future studies could record data from real candidates in driving exams or lessons, provided proper precautions regarding consent and data protection are taken.

It is also acknowledged that the current data proved to be specific to the vehicle used, as different vehicles have varying spring-damper characteristics, engine power, and therefore different acceleration capabilities. This can influence the accelerometer readings and should be considered when interpreting the results.

The use of sensors may contribute to increasing the efficiency of the driving test, and potentially provide valuable data for improving driver training programs. Data-based driving assessments may also prove useful to pre-license driver training and post-license driver monitoring.

Acknowledgments

We would like to thank the Dutch Central Office of Driving Certification (CBR) for making this research possible.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research is supported by Transitions and Behaviour grant [403.19.243] ("Towards Safe Mobility for All: A Data-Driven Approach"), provided by the Dutch Research Council (NWO).

ORCID

Tom Driessen http://orcid.org/0000-0003-1053-6020
Daniël Heikoop http://orcid.org/0000-0003-4379-072X
Dimitra Dodou http://orcid.org/0000-0002-9428-3261
Joost de Winter http://orcid.org/0000-0002-1281-8200

Data availability statement

Raw data and analysis scripts are available at https://doi.org/10.4121/3bb2f535-59ec-426c-b69a-e113810543b2

References

Alsharef, A., A. Albert, and S. Bhandari. 2022. "Safety Challenges Experienced by Driver License Examiners and Related Safety Measures." *Safety Science* 154:105808. https://doi.org/10.1016/j.ssci.2022.105808.

Alsharef, A., A. Albert, S. M. J. Uddin, N. B. Kittur, S. Chavan, and E. Jaselskis. 2021. "Work-Related Safety Incidents Among Driver License Examiners." *Safety Science* 140:105304. https://doi.org/10.1016/j.ssci.2021.105304.

Apple Inc. 2022. "Understanding Reference Frames and Device Attitude." https://web.archive.org/web/20220317180105/https://developer.apple.com/documentation/coremotion/getting_processed_device-motion_data/understanding_reference_frames_and_device_attitude.

Bagdadi, O., and A. Várhelyi. 2011. "Jerky Driving—An Indicator of Accident Proneness?" *Accident Analysis & Prevention* 43 (4): 1359–1363. https://doi.org/10.1016/j.aap.2011.02.009.

Bates, L., L. Anderson, D. Rodwell, and E. Blais. 2020. "A Qualitative Study of Young Drivers and Deterrence Based Road Policing." *Transportation Research Part F: Traffic Psychology and Behaviour* 71:110–118. https://doi.org/10.1016/j.trf.2020.04.003.

Baughan, C. J., N. P. Gregersen, M. Hendrix, and E. Keskinen. 2005. Towards European Standards for Testing. *Final Report*. Brussels: The International Commission for Driving Testing. https://www.cieca.eu/project/27 .

Bochkovskiy, A., C. Y. Wang, and H. Y. M. Liao. 2020. *YOLOv4: Optimal Speed and Accuracy of Object Detection. arXiv*. https://doi.org/10.48550/arXiv.2004.10934.

Cai, M., M. A. A. Yazdi, A. Mehdizadeh, Q. Hu, A. Vinel, K. Davis, H. Xian, F. M. Megahed, and S. E. Rigdon. 2021. "The Association Between Crashes and Safety-Critical Events: Synthesized Evidence from Crash Reports and Naturalistic Driving Data Among Commercial Truck Drivers." *Transportation Research Part C: Emerging Technologies* 126:103016. https://doi.org/10.1016/j.trc.2021.103016.

Carlos, M. R., L. C. González, J. Wahlström, G. Ramírez, F. Martínez, and G. Runger. 2020. "How Smartphone Accelerometers Reveal Aggressive Driving Behavior?—The Key Is the Representation." *IEEE Transactions on Intelligent Transportation Systems* 21 (8): 3377–3387. https://doi.org/10.1109/TITS.2019.2926639.

Chan, T. K., C. S. Chin, H. Chen, and X. Zhong. 2020. "A Comprehensive Review of Driver Behavior Analysis Utilizing Smartphones." *IEEE Transactions on Intelligent Transportation Systems* 21 (10): 4444–4475. https://doi.org/10.1109/TITS.2019.2940481.

Curry, A. E., K. B. Metzger, A. F. Williams, and B. C. Tefft. 2017. "Comparison of Older and Younger Novice Driver Crash Rates: Informing the Need for Extended Graduated Driver Licensing Restrictions." *Accident Analysis & Prevention* 108:66–73. https://doi.org/10.1016/j.aap.2017.08.015.

De Groot, S., J. C. F. De Winter, and M. Mulder. 2011. "Nonvestibular Motion Cueing in a Fixed-Base Driving Simulator: Effects on Driver Braking and Cornering Performance." *Presence* 20 (2): 117–142. https://doi.org/10.1162/pres_a_00039.

De Waard, D., and T. Rooijers. 1994. "An Experimental Study to Evaluate the Effectiveness of Different Methods and Intensities of Law Enforcement on Driving Speed on Motorways." *Accident Analysis & Prevention* 26 (6): 751–765. https://doi.org/10.1016/0001-4575(94)90052-3.

De Winter, J. C. F., S. De Groot, M. J. Van Loenhout, A. Van Leeuwen, P. Do, P. A. Wieringa, and M. Mulder. 2008. "Feedback on Mirror-Checking During Simulation-Based Driver Training." *Proceedings of the 27th European Annual Conference on Human Decision Making and Manual Control*, Delft, the Netherlands.

De Winter, J., J. Hoogmoed, J. Stapel, D. Dodou, and P. Bazilinskyy. 2023. "Predicting Perceived Risk of Traffic Scenes Using Computer Vision." *Transportation Research Part F: Traffic Psychology and Behaviour* 93:235–247. https://doi.org/10.1016/j.trf.2023.01.014.

Driessen, T., D. Dodou, D. De Waard, and J. C. F. De Winter. 2024. "Predicting Damage Incidents, Fines, and Fuel Consumption from Truck Driver Data: A Study from the Netherlands." *Transportation Research Record*. https://doi.org/10.1177/03611981231211897.

Driessen, T., A. Picco, D. Dodou, D. De Waard, and J. De Winter. 2021. "Driving Examiners' Views on Data-Driven Assessment of Test Candidates: An Interview Study." *Transportation Research Part F: Traffic Psychology and Behaviour* 83:60–79. https://doi.org/10.1016/j.trf.2021.09.021.

DS Automobiles. 2023. "DS Driver Attention Monitoring." https://www.dsautomobiles.co.uk/ds-experience/discover/news/ds-automobiles-anti-fatigue-technology.html.

Fell, J. C., K. Jones, E. Romano, and R. Voas. 2011. "An Evaluation of Graduated Driver Licensing Effects on Fatal Crash Involvements of Young Drivers in the United States." *Traffic Injury Prevention* 12 (5): 423–431. https://doi.org/10.1080/15389588.2011.588296.

Feng, F., S. Bao, J. R. Sayer, C. Flannagan, M. Manser, and R. Wunderlich. 2017. "Can Vehicle Longitudinal Jerk Be Used to Identify Aggressive Drivers? An Examination Using Naturalistic Driving Data." *Accident Analysis & Prevention* 104:125–136. https://doi.org/10.1016/j.aap.2017.04.012.

Fisher, D. L., N. E. Laurie, R. Glaser, K. Connerney, A. Pollatsek, S. A. Duffy, and J. Brock. 2002. "Use of a Fixed-Base Driving Simulator to Evaluate the Effects of Experience and PC-Based Risk Awareness Training on Drivers' Decisions." *Human Factors* 44 (2): 287–302. https://doi.org/10.1518/0018720024497853.

Foxe, K. 2020. More Than 50 Incidents of Aggression by Driving Test Candidates Reported. https://www.irishexaminer.com/news/arid-40195446.html.

Harwood, D. W. 1992. "Traffic and Vehicle Operating Characteristics." *Traffic Engineering Handbook*, edited by J. L. Pline, 29–58. 4th ed. Englewood Cliffs, NJ: Prentice Hall.

Helman, S., W. Vlakveld, B. Fildes, J. Oxley, K. Fernández-Medina, and J. Weekley. 2017. Study on Driver Training, Testing and Medical Fitness. *Final report*. https://doi.org/10.2832/446103.

Hunter, M., E. Saldivar-Carranza, J. Desai, J. K. Mathew, H. Li, and D. M. Bullock. 2021. "A Proactive Approach to Evaluating Intersection Safety Using Hard-Braking Data." *Journal of Big Data Analytics in Transportation* 3 (2): 81–94. https://doi.org/10.1007/s42421-021-00039-y.

Itkonen, T. H., J. Pekkanen, O. Lappi, I. Kosonen, T. Luttinen, H. Summala, and F. Lucidi. 2017. "Trade-Off Between Jerk and Time Headway as an Indicator of Driving Style." *PLOS ONE* 12 (10): e0185856. https://doi.org/10.1371/journal.pone.0185856.

Johnson, D. A., and M. M. Trivedi. 2011. "Driving Style Recognition Using a Smartphone As a Sensor Platform." *Proceedings of the 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 1609–1615. Washington, DC. https://doi.org/10.1109/ITSC.2011.6083078.

Khorram, B., A. E. Af Wåhlberg, and A. Tavakoli Kashani. 2020. "Longitudinal Jerk and Celeration as Measures of Safety in Bus Rapid Transit Drivers in

Tehran." *Theoretical Issues in Ergonomics Science* 21 (5): 577–594. https://doi.org/10.1080/1463922X.2020.1719228.

Kim, H., S. Martin, A. Tawari, T. Misu, and J. L. Gabbard. 2020. "Toward Real-Time Estimation of Driver Situation Awareness: An Eye-Tracking Approach Based on Moving Objects of Interest." *Proceedings of the 2020 IEEE Intelligent Vehicles Symposium*, 1035–1041. Las Vegas, NV. https://doi.org/10.1109/IV47402.2020.9304770.

Lajunen, T., M. J. M. Sullman, and E. Gaygısız. 2022. "Self-Assessed Driving Skills and Risky Driver Behaviour Among Young Drivers: A Cross-Sectional Study." *Frontiers in Psychology* 13:840269. https://doi.org/10.3389/fpsyg.2022.840269.

Lin, T.-Y., M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. 2014. "Microsoft COCO: Common Objects in Context." *Proceedings of the European Conference on Computer Vision*, 740–755. Zurich, Switzerland. https://doi.org/10.1007/978-3-319-10602-1_48.

Mapbox. 2023. "Map Matching API." https://docs.mapbox.com/api/navigation/map-matching.

Marafie, Z., K.-J. Lin, D. Wang, H. Lyu, Y. Liu, Y. Meng, and J. Ma. 2021. "AutoCoach: An Intelligent Driver Behavior Feedback Agent with Personality-Based Driver Models." *Electronics* 10 (11): 1361. https://doi.org/10.3390/electronics10111361.

Mathworks. 2021. "Sensor Data Collection with MATLAB Mobile." https://nl.mathworks.com/help/matlabmobile/ug/sensor-data-collection-with-matlab-mobile.html.

Ma, Y.-L., X. Zhu, X. Hu, and Y.-C. Chiu. 2018. "The Use of Context-Sensitive Insurance Telematics Data in Auto Insurance Rate Making." *Transportation Research Part A: Policy and Practice* 113:243–258. https://doi.org/10.1016/j.tra.2018.04.013.

Melman, T., D. Abbink, X. Mouton, A. Tapus, and J. De Winter. 2021. "Multivariate and Location-Specific Correlates of Fuel Consumption: A Test Track Study." *Transportation Research Part D: Transport & Environment* 92:102627. https://doi.org/10.1016/j.trd.2020.102627.

Mueller, A. S., and J. B. Cicchino. 2022. "Teen Driver Crashes Potentially Preventable by Crash Avoidance Features and Teen-Driver-Specific Safety Technologies." *Journal of Safety Research* 81:305–312. https://doi.org/10.1016/j.jsr.2022.03.007.

Musicant, O., and T. Lotan. 2016. "Can Novice Drivers Be Motivated to Use a Smartphone Based App That Monitors Their Behavior?" *Transportation Research Part F: Traffic Psychology and Behaviour* 42:544–557. https://doi.org/10.1016/j.trf.2015.10.023.

Nambi, A. U., I. Mehta, A. Ghosh, V. Lingam, and V. N. Padmanabhan. 2019. "ALT: Towards Automating Driver License Testing Using Smartphones." *Proceedings of the 17th Conference on Embedded Networked Sensor Systems*, 29–42. New York, NY. https://doi.org/10.1145/3356250.3360037.

Othman, W., A. Kashevnik, B. Hamoud, and N. Shilov. 2022. "DriverSVT: Smartphone-Measured Vehicle Telemetry Data for Driver State Identification." *Data* 7 (12): 181. https://doi.org/10.3390/data7120181.

Parmet, Y., A. Borowsky, O. Yona, and T. Oron-Gilad. 2015. "Driving Speed of Young Novice and Experienced Drivers in Simulated Hazard Anticipation Scenes." *Human Factors: The Journal of the Human Factors & Ergonomics Society* 57 (2): 311–328. https://doi.org/10.1177/0018720814548220.

Pedley, M. 2013. Tilt Sensing Using a Three-Axis Accelerometer. *Report No. AN3461*. https://www.nxp.com/docs/en/application-note/AN3461.pdf.

Poirier, B., E. Blais, and C. Faubert. 2018. "Graduated Driver Licensing and Differential Deterrence: The Effect of License Type on Intentions to Violate Road Rules." *Accident Analysis & Prevention* 110:62–70. https://doi.org/10.1016/j.aap.2017.10.001.

Qin, L., Y. Shi, Y. He, J. Zhang, X. Zhang, Y. Li, T. Deng, and H. Yan. 2022. "ID-YOLO: Real-Time Salient Object Detection Based on the Driver's Fixation Region." *IEEE Transactions on Intelligent Transportation Systems* 23 (9): 15898–15908. https://doi.org/10.1109/TITS.2022.3146271.

Rezaei, M., M. Azarmi, and F. M. P. Mir. 2021. "Traffic-Net: 3D Traffic Monitoring Using a Single Camera." arXiv. https://doi.org/10.48550/arXiv.2109.09165.

Roemer, E. 2021. *Van Rijles naar Rijonderwijs. Advies Verbetering Autorijscholenbranche [From Driving Lessons to Driving Education. Advice for Improving the Driving School Industry]*. The Hague, The Netherlands: Ministry of Infrastructure and Water Management. https://www.rijksoverheid.nl/documenten/rapporten/2021/04/14/bijlage-1-van-rijles-naar-rijonderwijs-advies-verbeteren-autorijscholenbranche.

Rolison, J. J., and S. Moutari. 2020. "Combinations of Factors Contribute to Young Driver Crashes." *Journal of Safety Research* 73:171–177. https://doi.org/10.1016/j.jsr.2020.02.017.

sbairagy-MW. 2021. "Pretrained-yolo-v4." https://github.com/matlab-deep-learning/pretrained-yolo-v4

Senserrick, T., and N. Haworth. 2005. Review of Literature Regarding National and International Young Driver Training, Licensing and Regulatory Systems. *Report No. 239*. Monash University. https://www.monash.edu/muarc/archive/our-publications/reports/muarc239.

Shanly, C., M. Ieti, I. Warren, and J. Sun. 2018. "BackPocketDriver-A Mobile App to Enhance Safe Driving for Youth." *Proceedings of the International Conference on Software Engineering and Knowledge Engineering*, 246–245. San Francisco Bay, CA. https://doi.org/10.18293/SEKE2018-011.

Stipancic, J., L. Miranda-Moreno, and N. Saunier. 2016. "The Who and Where of Road Safety: Extracting Surrogate Indicators from Smartphone-Collected GPS Data in Urban Environments." *Transportation Research Board 95th Annual Meeting*, Washington, DC. https://www.n.saunier.free.fr/saunier/stock/stipancic16who-where.pdf.

Stipancic, J., L. Miranda-Moreno, and N. Saunier. 2018. "Vehicle Manoeuvers as Surrogate Safety Measures: Extracting Data from the GPS-Enabled Smartphones of Regular Drivers." *Accident Analysis & Prevention* 115:160–169. https://doi.org/10.1016/j.aap.2018.03.005.

Takeda, K., C. Miyajima, T. Suzuki, P. Angkititrakul, K. Kurumida, Y. Kuroyanagi, H. Ishikawa, et al. 2012. "Self-Coaching System Based on Recorded Driving Data: Learning from One's Experiences." *IEEE Transactions on Intelligent Transportation Systems* 13 (4): 1821–1831. https://doi.org/10.1109/TITS.2012.2205917.

Tay, R. 2005. "The Effectiveness of Enforcement and Publicity Campaigns on Serious Crashes Involving Young Male Drivers: Are Drink Driving and Speeding Similar?" *Accident Analysis & Prevention* 37 (5): 922–929. https://doi.org/10.1016/j.aap.2005.04.010.

Telemetry Overlay S.L. 2022. "GPS & Data Visualization." https://goprotelemetryextractor.com

UK Government. 2022. "Top 10 Reasons for Failing the Driving Test in Great Britain." https://www.gov.uk/government/publications/top-10-reasons-for-failing-the-driving-test/top-10-reasons-for-failing-the-driving-test-in-great-britain.

Weast, R. A., and S. S. Monfort. 2021. "Characteristics of Vehicles Driven by Teens and Adults Killed in Crashes, 2013–2017." *Journal of Safety Research* 77:263–267. https://doi.org/10.1016/j.jsr.2021.03.010.

Williams, A. F. 2017. "Graduated Driver Licensing (GDL) in the United States in 2016: A Literature Review and Commentary." *Journal of Safety Research* 63:29–41. https://doi.org/10.1016/j.jsr.2017.08.010.

Appendix: Example rows of raw data collected by iPhone X and GoPro Max

Tables A1–A3 show example rows of data recorded.

Table A1. Example of phone acceleration data (measurement frequency: 10 Hz).

Timestamp	X (m/s^2)	Y (m/s^2)	Z (m/s^2)
04-04-2022 10:13:38.186	−0.236	6.940	7.216
04-04-2022 10:13:38.287	−0.626	6.955	7.348
04-04-2022 10:13:38.387	0.126	7.619	6.059
04-04-2022 10:13:38.488	0.869	7.293	7.691
04-04-2022 10:13:38.589	1.997	6.670	5.816
:	:	:	:
04-04-2022 10:46:42.163	−1.219	4.693	8.316
04-04-2022 10:46:42.264	−1.544	4.919	8.676
04-04-2022 10:46:42.364	2.094	6.665	6.603
04-04-2022 10:46:42.465	−1.113	6.537	7.096
04-04-2022 10:46:42.566	−1.338	5.267	8.178

Table A2. Example of GoPro acceleration data (measurement frequency: approximately 200 Hz, except for the first few samples).

Timestamp	X (m/s^2)	Y (m/s^2)	Z (m/s^2)
2022-04-04 10:13:29.024	−1.012	−1.122	−9.827
2022-04-04 10:13:29.036	−1.405	−1.275	−9.508
2022-04-04 10:13:29.048	−1.309	−1.048	−9.340
2022-04-04 10:13:29.061	−0.815	−0.686	−9.892
2022-04-04 10:13:29.073	−0.206	−0.341	−10.251
:	:	:	:
2022-04-04 10:46:57.012	−0.180	−0.638	−9.765
2022-04-04 10:46:57.017	−0.751	−1.031	−9.892
2022-04-04 10:46:57.022	−1.290	−1.185	−9.731
2022-04-04 10:46:57.028	−1.393	−1.048	−9.326
2022-04-04 10:46:57.033	−1.028	−0.847	−9.925

Table A3. Example of GoPro GPS data (the measurement frequency was fluctuating but averaged approximately 17 Hz).

Timestamp	Latitude	Longitude	Altitude	2D Speed
2022-04-04 10:37:05.715	52.144	5.388	−2.477	11.105
2022-04-04 10:37:05.773	52.144	5.388	−2.516	11.129
2022-04-04 10:37:05.831	52.144	5.388	−2.513	11.226
2022-04-04 10:37:05.889	52.144	5.388	−2.502	11.219
2022-04-04 10:37:05.947	52.144	5.388	−2.532	11.188
:	:	:	:	:
2022-04-04 10:46:57.108	52.144	5.423	−4.613	0.010
2022-04-04 10:46:57.189	52.144	5.423	−4.619	0.070
2022-04-04 10:46:57.269	52.144	5.423	−4.625	0.050
2022-04-04 10:46:57.349	52.144	5.423	−4.661	0.040
2022-04-04 10:46:57.430	52.144	5.424	−4.652	0.020