Distinguishing Level 0 and Level 2 driving mode via spectral analysis of lateral position: Insights from a highway Naturalistic Driving Study

Nikolaos Panagiotis Trantas*[1], Konstantinos Mattas[2], Riccardo Dona[2], Akos Kriston[2], Biagio Ciuffo[2], Constantinos Antoniou[1]

[1] Chair of Transportation Systems Engineering, Technical University of Munich, Germany

[2] Project Officer, Directorate for Energy Transport and Climate, European Commission – Joint Research Centre, Via E. Fermi, 2749, 21027 Ispra (VA), Italy

SHORT SUMMARY

The development of Level 2 driving systems complicates the interpretation of Naturalistic Driving Study (NDS) data, as behavioral modeling requires accurate identification of the operational mode. While longitudinal control has already been studied, lateral behavior remains underexplored, particularly in the frequency domain. This study proposes a classification framework to distinguish between SAE Level 0 and Level 2 driving using the Power Spectral Density (PSD) of the vehicle's lateral position. High-precision trajectory data (10 Hz) from six instrumented vehicles were collected on the I-24 MOTION testbed, and the signals were split into short driving episodes. We extract five spectral metrics, band power, slope, centroid, band power ratio, and entropy, to capture distinct control signatures. Statistical analysis confirms significant differences between driving modes across all examined time horizons (15-90 s). A logistic regression model achieves a peak accuracy of 92% at a 60-second observation window, identifying this as the optimal duration. These results demonstrate that frequency-domain methods effectively separate machine control from human steering, providing a foundational tool for large-scale safety analysis and automated dataset labeling.

Keywords: Driving behavior, Automated vehicles, Human driving, Driver type identification, Spectral analysis, Vehicle trajectories

1. INTRODUCTION

The gradual deployment of partially automated vehicles is fundamentally reshaping the road environment. While SAE Level 1 technologies have been standard for over a decade, the recent adoption of Driver Control Assistance Systems (DCAS) under UN Regulation No. 171 (UNECE, 2024) is expected to accelerate the global integration of Level 2 systems. In the context of Naturalistic Driving Studies (NDS), this mixed-traffic reality complicates data interpretation. Identifying the operational modes of both ego and neighboring vehicles has become a prerequisite for reliable behavioral modeling; without knowing "who" is in control, the validity of safety and traffic flow analyses is called into question.

Effective vehicle monitoring requires distinguishing between manual and automated control through externally observable behavior, a task with significant "forensic" value for accident reconstruction and large-scale dataset labeling. Behavioral research has already focused on the longitudinal dimension of driving (Marín et al., 2025, Liu et al., 2015). However, longitudinal control is frequently governed by exogenous factors, such as the behavior of a lead vehicle or traffic density, which can mask the underlying control logic. In contrast, lateral lane positioning offers

a more stable signal with high explanatory power, especially regarding inattention and control authority (Papantoniou, 2018), with the added benefit of being readily measurable via standard NDS sensor suites. Despite its potential, the lateral dimension remains underexplored, particularly regarding its latent control signatures.

A critical gap exists in the methodology used to analyze these control signals. Traditionally, frequency-domain (f-domain) methods in road safety research have been applied in biometric or physiological signals, such as EEG (Gurudath & Riley, 2014). Frequency-domain analysis is complementary to time-domain approaches; with respect to vehicle dynamics it can reveal systematic control "fingerprints" and periodicities, such as the specific oscillation patterns of a Level 2 centering algorithm, that are often obscured in the time domain. Nevertheless, regarding observable driving behavior, it should be noted that driving is a closed-loop control process in which the system output (vehicle trajectory) is not a "pure" representation of the driver's input. Rather, the observed signal is a diluted representation incorporating vehicle dynamics and measurement noise, such as GNSS inaccuracies. Any forensic inference of behavior must therefore treat the trajectory as the product of a combined human-vehicle-measurement system, and not as equivalent of human control input.

This article proposes a method to identify SAE Level 0 and Level 2 driving modes, by analyzing the driving signal in the f-domain. We develop suitable metrics to enumerate the calculated spectra and infer behavioral signatures. The primary contribution of this study is a classification framework to predict the automation mode based on short driving episodes recorded in a Naturalistic Driving Study. Furthermore, we evaluate how the observation duration influences prediction performance to determine the optimal window for driver type identification. The remainder of this paper is structured as follows: Section 2 details the methodology for spectral feature extraction and model specification. Section 3 presents the dataset and discusses the statistical significance of the identified metrics and the classification results. Finally, Section 4 summarizes the findings and outlines future research directions.

2. METHODOLOGY

Spectral analysis

To isolate behavioral control signatures, lateral position signals across fixed-duration driving episodes are first detrended to eliminate linear drift and center the observations. The Power Spectral Density (PSD) is estimated using the Welch method (Zhao & He, 2013a) with a Hanning window of 128 observations and a 50% overlap. This window size corresponds to the largest power-of-2 compatible with the shortest examined observation horizon, ensuring parameter consistency across all durations. Given the Nyquist limit, the resulting PSD spans the 0-5 Hz band with a frequency resolution of 0.078 Hz.

Five spectral metrics are extracted to quantify the behavioral signatures of human and automated driving.

1. Band power, in the frequency band 0-1 Hz: the square root of the area under the PSD (calculated via Simpson's rule), representing the time-domain variance or vehicle "sway" (Eq. 1).
2. Spectral slope, in the frequency band 0.2-1 Hz: the gradient of the log-log PSD, indicating the rate at which control power decays from low to high frequencies (Eq. 2).
3. Spectral centroid: the "center of mass" of the power distribution, encapsulating the spectral shape into a single descriptive value (Eq. 3).

4. Band power ratio: the ratio of high-frequency (1-2 Hz) to low-frequency (0-1 Hz) power, capturing the intensity of small steering corrections relative to larger ones (Eq. 4).
5. Spectral entropy: a measure of spectral complexity, for a PSD with N frequency bins, bounded between 0 (pure sinusoid) and 1 (white noise) (Eq. 5).

$$BP = \left(\int_{f_n}^{f_m} PSD(f) df \right)^{1/2} \tag{1}$$

$$PSD(f) \propto f^{-a} \;\rightarrow\; log_{10}\big(PSD(f)\big) = -a \cdot log_{10}(f) + c \tag{2}$$

$$SC = \frac{\sum_{i=n}^{m} f_i \cdot PSD(f_i)}{\sum_{i=n}^{m} PSD(f_i)} \tag{3}$$

$$BPR = \frac{BP_{high\ band}{}^2}{BP_{low\ band}{}^2} \tag{4}$$

$$SE = \frac{-\sum_{i=1}^{N} p_i ln(p_i)}{ln(N)} \;, with\ p_i = \frac{PSD_i}{\sum_{j=1}^{N} PSD_j} \tag{5}$$

Statistical analysis

Metric distributions are evaluated for normality (Shapiro-Wilk, α=0.05) and equal variances (Levene's test, α=0.05). Differences in means between Level 0 and Level 2 are assessed via Student's t-test, Welch's t-test, or Mann-Whitney U test, depending on the distribution properties. The effect sizes of the distributions are quantified using Cohen's d (for normal distributions) or Cliff's delta (for non-normal distributions) (Romano et al., 2006, October).

Model specification

For the classification task, a balanced dataset of driving episodes is randomly partitioned into training (75%) and test (25%) subsets. A Logistic Regression model (Eq. 6) is specified using the z-score standardized spectral metrics as predictors. Predictor selection prioritizes uncorrelated features to ensure model parsimony, while the significance of the metrics is evaluated for different driving episode durations. Final predictive power is evaluated exclusively on the test subset using two performance metrics: accuracy and the area under the ROC Curve (AUC), with significance determined by 95% confidence intervals calculated via the Hanley & McNeil method (Hanley & McNeil, 1982).

$$P(Y = 1|\boldsymbol{X}) = \frac{1}{1 + e^{-\boldsymbol{\beta X}}} \tag{6}$$

3. RESULTS AND DISCUSSION

Data collection and treatment

In Autumn 2024, the Joint Research Centre and Vanderbilt University's Institute for Software Integrated Systems conducted a week-long experimental campaign on the I-24 MOTION testbed in Nashville, Tennessee, using six instrumented vehicles. This four-mile interstate segment is equipped with 276 ultra-high-definition cameras, enabling the creation of a comprehensive digital twin of the traffic flow (Gloudemans et al., 2023).

The fleet featured advanced SAE Level 2 systems, capable of taking over both longitudinal and lateral control, and performing assisted lane changes. Trials were conducted during morning and evening peak hours, in heavy congestion and free-flow conditions. These tests aimed to investigate stability phenomena of the longitudinal controller, assisted lane changing characteristics, and the collective impact of automated vehicles on surrounding traffic. Manual driving sessions were also recorded, establishing a baseline for comparison.

Each vehicle was manned by a safety driver and at least one researcher. For data acquisition, the technical service provider IDIADA installed a GNSS/IMU solution for high-precision trajectory tracking at a 10 Hz recording frequency, alongside three cameras monitoring the forward view, rear view, and instrument cluster.

This study focuses exclusively on the data from these six vehicles. Automation levels for each lap were manually annotated by the onboard researchers and verified via internal camera recordings. While this analysis relies on GNSS/IMU trajectories rather than external I-24 MOTION camera data, the coordinates were projected onto the I-24 MOTION curvilinear system. This projection enables the precise identification of lane positions and maneuvers. For the purposes of this work, only segments of steady, straight-line travel are isolated, effectively filtering out lane changes and curved road segments. These segments are partitioned into successive, non-overlapping driving episodes with durations of 15, 30, 45, 60, 75 and 90 seconds. Using non-overlapping windows ensures statistical independence of the samples, mitigating autocorrelation bias.

Results

Figure 1 depicts multiple PSDs in overlay, with each line representing one driving episode of the same duration. When plotted in log-log scale, the raw spectra exhibit a clear $1/f^{a}$ decay at the lower frequencies. Also at the lower frequencies, the manual driving signals appear to have more power while exhibiting less variance and being more tightly grouped together, hinting on a more homogeneous control authority of the human drivers.

Table 1 summarizes the statistical analysis. To interpret the effect sizes, the standard benchmarks are adopted: for Cohen's d, values of 0.2, 0.5 and 0.8 represent small, medium and large effects respectively; for Cliff's delta, thresholds of 0.147, 0.33 and 0.474 denote negligible, small, medium and large distances (Romano et al., 2006). These thresholds, being arbitrary and domain-specific, provide a first indication of distribution separation and, most importantly, serve as a reference to quantify the relative change across different examined durations.

Two interesting trends are evident. First, the difference in the means is statistically significant for all metrics and durations. Second, effect sizes consistently increase with duration, albeit with a subtle stabilization from 75 to 90 s. This suggests that while spectral signatures are present in

short windows, they solidify as the observation horizon extends, providing greater explanatory power.

Figure 1: Magnified visualization of the PSD in logarithmic axes

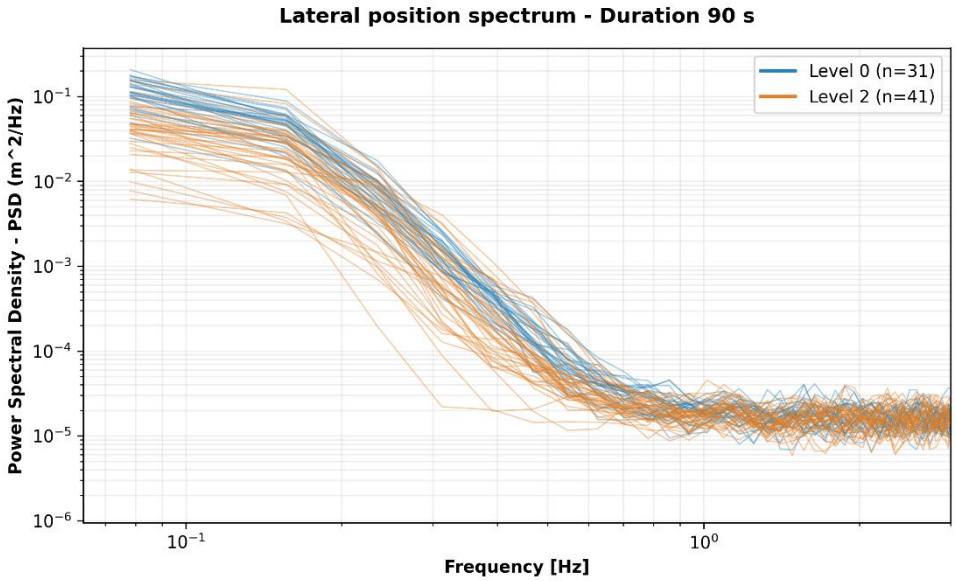


Table 1: Statistical testing of the spectral metrics

Metric	Sample (L0, L2)	Test	L0 Mean ± SE	L2 Mean ± SE	p-value	Eff. size
Band power						
15 s	208, 284	Mann-Whit	0.081 ± 0.003	0.056 ± 0.002	0.0000	delta = 0.37
30 s	102, 137	Mann-Whit	0.088 ± 0.003	0.062 ± 0.003	0.0000	delta = 0.48
45 s	64, 89	T-Test	0.090 ± 0.003	0.063 ± 0.003	0.0000	d = 1.06
60 s	49, 65	T-Test	0.091 ± 0.003	0.065 ± 0.003	0.0000	d = 1.22
75 s	38, 49	Mann-Whit	0.089 ± 0.003	0.062 ± 0.003	0.0000	delta = 0.71
90 s	31, 41	T-Test	0.091 ± 0.003	0.064 ± 0.003	0.0000	d = 1.34
Spectral slope						
15 s	208, 284	Welch-T	4.069 ± 0.076	3.559 ± 0.078	0.0000	d = 0.42
30 s	102, 137	Welch-T	4.232 ± 0.069	3.581 ± 0.087	0.0000	d = 0.73
45 s	64, 89	Welch-T	4.247 ± 0.068	3.678 ± 0.095	0.0000	d = 0.74
60 s	49, 65	Mann-Whit	4.196 ± 0.071	3.668 ± 0.104	0.0003	delta = 0.40
75 s	38, 49	Welch-T	4.241 ± 0.074	3.716 ± 0.110	0.0002	d = 0.80
90 s	31, 41	Welch-T	4.289 ± 0.076	3.778 ± 0.125	0.0009	d = 0.77
Spectral centroid						
15 s	208, 284	Mann-Whit	0.122 ± 0.002	0.128 ± 0.002	0.0007	delta = -0.18
30 s	102, 137	Mann-Whit	0.114 ± 0.001	0.122 ± 0.001	0.0003	delta = -0.27
45 s	64, 89	Mann-Whit	0.111 ± 0.001	0.120 ± 0.002	0.0000	delta = -0.40
60 s	49, 65	Mann-Whit	0.110 ± 0.001	0.119 ± 0.002	0.0002	delta = -0.41
75 s	38, 49	Mann-Whit	0.111 ± 0.001	0.120 ± 0.002	0.0004	delta = -0.44
90 s	31, 41	Mann-Whit	0.110 ± 0.001	0.120 ± 0.002	0.0008	delta = -0.46
Band power ratio						
15 s	208, 284	Mann-Whit	0.006 ± 0.001	0.015 ± 0.002	0.0000	delta = -0.37
30 s	102, 137	Mann-Whit	0.003 ± 0.000	0.010 ± 0.001	0.0000	delta = -0.46

45 s	64, 89	Mann-Whit	0.003 ± 0.000	0.008 ± 0.001	0.0000	delta = -0.54
60 s	49, 65	Mann-Whit	0.002 ± 0.000	0.006 ± 0.001	0.0000	delta = -0.57
75 s	38, 49	Mann-Whit	0.002 ± 0.000	0.006 ± 0.001	0.0000	delta = -0.70
90 s	31, 41	Mann-Whit	0.002 ± 0.000	0.006 ± 0.001	0.0000	delta = -0.66
Spectral entropy						
15 s	208, 284	Mann-Whit	0.267 ± 0.005	0.309 ± 0.006	0.0000	delta = -0.28
30 s	102, 137	Mann-Whit	0.272 ± 0.005	0.307 ± 0.006	0.0001	delta = -0.30
45 s	64, 89	Mann-Whit	0.267 ± 0.004	0.303 ± 0.006	0.0002	delta = -0.35
60 s	49, 65	Mann-Whit	0.264 ± 0.004	0.297 ± 0.006	0.0002	delta = -0.40
75 s	38, 49	Mann-Whit	0.267 ± 0.004	0.301 ± 0.006	0.0001	delta = -0.48
90 s	31, 41	Mann-Whit	0.267 ± 0.004	0.300 ± 0.007	0.0006	delta = -0.48

The model specification results are presented in Table 2. To avoid multicollinearity, "Band power ratio" is excluded in favor of "Band power", which demonstrates higher predictive power. It is interesting that some metrics have alternating significance as the duration increases; for instance, spectral entropy is very significant for very short driving episodes, whereas spectral slope is more significant for the middle ones. This implies that high-frequency "jitter" (entropy) is a short-term indicator of automation, while the systematic "decay" of control (slope) requires a longer window to be accurately captured in the data. At the same time, spectral centroid is a relatively consistent predictor for the automation mode.

Table 2: Model specification for the logistic regression classifier

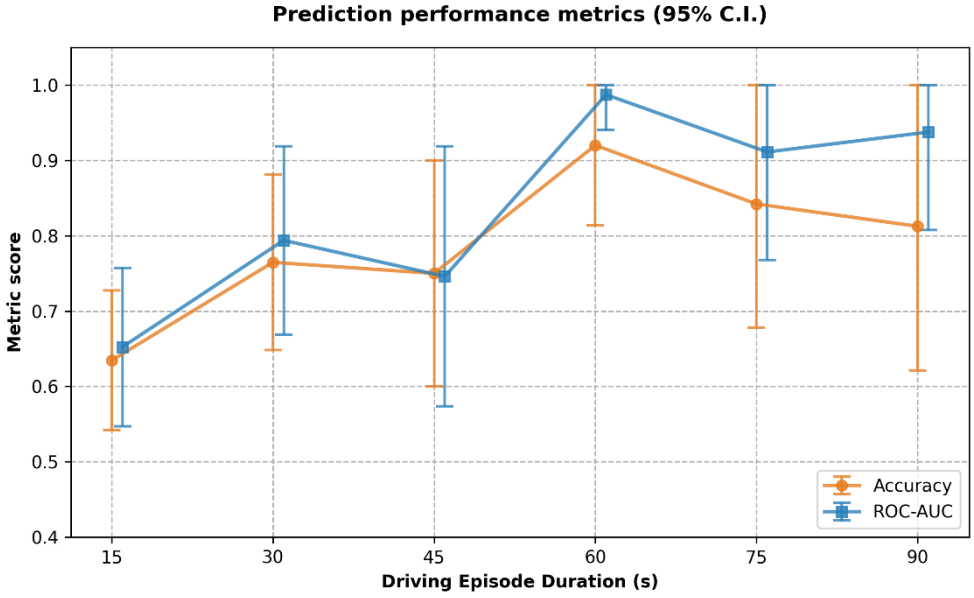
Feature	Coeff.	Std. Err.	p-value
Duration 15 s (Train N=312, Test N=104)			
Accuracy: 0.63 (95% CI: 0.54-0.73)			
ROC-AUC: 0.65 (95% CI: 0.55-0.76)			
Spectral slope	-0.230	0.145	0.1134
Spectral centroid	-0.279	0.186	0.1340
Spectral entropy	0.615	0.220	0.0051
Band power	-0.449	0.178	0.0115
Duration 30 s (Train N=153, Test N=51)			
Accuracy: 0.76 (95% CI: 0.65-0.88)			
ROC-AUC: 0.79 (95% CI: 0.67-0.92)			
Spectral slope	-0.903	0.255	0.0004
Spectral centroid	0.915	0.388	0.0182
Spectral entropy	-0.315	0.442	0.4758
Band power	-0.348	0.301	0.2478
Duration 45 s (Train N=96, Test N=32)			
Accuracy: 0.75 (95% CI: 0.60-0.90)			
ROC-AUC: 0.75 (95% CI: 0.57-0.92)			
Spectral slope	-1.600	0.469	0.0006
Spectral centroid	2.408	0.706	0.0007
Spectral entropy	-1.527	0.781	0.0506
Band power	-1.223	0.549	0.0260
Duration 60 s (Train N=73, Test N=25)			
Accuracy: 0.92 (95% CI: 0.81-0.99 *)			
ROC-AUC: 0.99 (95% CI: 0.94-0.99 *)			
Spectral slope	-1.228	0.521	0.0184
Spectral centroid	2.667	0.930	0.0041
Spectral entropy	-1.008	0.831	0.2251
Band power	-1.287	0.561	0.0217

Duration 75 s (Train N=57, Test N=19)			
Accuracy: 0.84 (95% CI: 0.68-0.99 *)			
ROC-AUC: 0.91 (95% CI: 0.77-0.99 *)			
Spectral slope	-1.667	0.638	0.0090
Spectral centroid	3.267	1.269	0.0101
Spectral entropy	-0.855	1.165	0.4630
Band power	-1.188	0.690	0.0850
Duration 90 s (Train N=46, Test N=16)			
Accuracy: 0.81 (95% CI: 0.62-0.99 *)			
ROC-AUC: 0.94 (95% CI: 0.81-0.99 *)			
Spectral slope	-0.579	0.717	0.4196
Spectral centroid	3.934	1.695	0.0203
Spectral entropy	-2.675	1.535	0.0814
Band power	-2.412	1.087	0.0265

** The interval is truncated by the physical limits of the metric.*

Predictive performance, visualized in Figure 2, improves significantly with duration. Accuracy reaches a peak value of 92% at the 60 s horizon. While accuracy and AUC remain high at 75 s and 90 s, the broadening of the 95% confidence intervals is likely an artifact of the diminishing size of the test subset (19 and 16 observations at the 75 and 90 s respectively). Concurrently, the plateau indicates that the model's predictive power stabilizes, in contrast to the abrupt improvement at 60 s. Consequently, the 60 s window appears to be an optimal "sweet spot" that balances high discriminative power with statistical validity.

Figure 2: Model performance for all the examined durations



4. CONCLUSIONS

This study demonstrates that frequency-domain analysis of lateral vehicle position provides a robust framework for distinguishing between manual (Level 0) and partially automated (Level 2) driving. By extracting spectral features that capture the latent behavioral signatures of control authority, we successfully developed a classification model that achieves 92% accuracy within a

60-second observation horizon. This duration offers a benchmark for the automated labeling of Naturalistic Driving Study (NDS) data, addressing a critical gap in behavioral modeling within mixed-traffic environments.

However, the analysis in this study is restricted to straight-line driving, omitting more complex turning maneuvers. Future research will extend this methodology to include road curves and lane changes, where the divergence between human steering entropy and algorithmic trajectory planning is expected to be more pronounced. Furthermore, while the 60-second window proved optimal for our current dataset, the broader confidence intervals at 75 and 90 seconds underscore the need for larger-scale validation to stabilize performance metrics at longer horizons. Additionally, more classification models will be assessed to improve predictive power at shorter time horizons. Ultimately, this work establishes a non-intrusive, frequency-based methodology for identifying automation operational modes, providing researchers with a complete tool for driving behavior identification and large-scale road safety analysis.

ACKNOWLEDGEMENTS

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 571155448.

REFERENCES

Gloudemans, D., Wang, Y., Ji, J., Zachár, G., Barbour, W., Hall, E., Cebelak, M., Smith, L., Work, D.B., 2023. I-24 MOTION: An instrument for freeway traffic science. Transportation Research Part C: Emerging Technologies 155, 104311. https://doi.org/10.1016/j.trc.2023.104311

Gurudath, N., & Riley, H. B. (2014). Drowsy driving detection by EEG analysis using wavelet transform and K-means clustering. *Procedia Computer Science*, *34*, 400-409.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, *143*(1), 29-36.

Liu, Z., Ivanco, A., & Filipi, Z. (2015). Quantification of drive cycle's rapid speed fluctuations using Fourier analysis. *SAE International Journal of Alternative Powertrains*, *4*(1), 170-177.

Marín, A. L., Martínez-Plumed, F., Ramírez-Quintana, M. J., Mattas, K., Fontaras, G., Kouvelas, A., & Makridis, M. A. (2025). Human or machine: A novel deep learning framework for autonomous driver identification based on vehicle trajectories. *IEEE Transactions on Intelligent Transportation Systems*.

Papantoniou, P. (2018). Structural equation model analysis for the evaluation of overall driving performance: A driving simulator study focusing on driver distraction. Traffic injury prevention, 19(3), 317-325.

Romano, J., Kromrey, J. D., Coraggio, J., Skowronek, J., & Devine, L. (2006, October). Exploring methods for evaluating group differences on the NSSE and other surveys: Are the t-test and Cohen'sd indices the most appropriate choices. In *annual meeting of the Southern Association for Institutional Research* (Vol. 14). Citeseer.

United Nations Economic Commission for Europe. (2024). *Proposal for a new UN Regulation on uniform provisions concerning the approval of vehicles with regard to Driver Control Assistance Systems (DCAS)* (ECE/TRANS/WP.29/2024/37). https://unece.org/transport/documents/2024/02/working-documents/grva-proposal-new-un-regulation-uniform-provisions

Zhao, L., & He, Y. (2013). Power spectrum estimation of the welch method based on imagery EEG. *Applied Mechanics and Materials*, *278*, 1260-1264.