Driver Classification Using Self-reported, Psychophysiological, and Performance Metrics within a Simulated Environment (Extended Abstract)

Vishal C. Kummetha[1], Umair Durrani, Justin Mason, Sisinnio Concas, Alexandra Kondyli

Center for Urban Transportation Research (CUTR), University of South Florida, 4202 E. Fowler Ave, Tampa FL, kummetha@usf.edu (Vishal Kummetha, Ph.D.)
Department of Civil and Environmental Engineering, University of Windsor, 401 Sunset Avenue, Windsor ON, durraniu@uwindsor.ca (Umair Durrani)
Department of Occupational Therapy, University of Florida, 1225 Center Drive, Gainesville FL, justinmason@php.ufl.edu (Justin Mason, Ph.D.)
Center for Urban Transportation Research (CUTR), University of South Florida, 4202 E. Fowler Ave, Tampa FL, concas@usf.edu (Sisinnio Concas, Ph.D.)
Department of Civil, Environmental, and Architectural Engineering, University of Kansas, 1530 W 15th St, Lawrence KS, akondyli@ku.edu (Alexandra Kondyli, Ph.D.)

Abstract

Driver classification provides an efficient approach to isolating unique traits associated with specific driver types under various driving conditions. Several past studies use classification to identify behavior and driving styles, however, very few studies employ both measurable physiological changes and environmental factors. This study looked to address the shortcomings in driver classification research using a data-driven approach to assess driving tasks performed under varying mental workloads. Psychophysiological and driving performance changes experienced by drivers when engaged in simulated tasks of varying difficulty were coupled with machine-learning techniques to provide a more accurate estimate of the ground truth for behavioral classification. A simulator study consisting of six tasks was carefully designed to incrementally vary complexity between individual tasks. Ninety drivers were recruited to participate in both the subjective and driving components of this research. Positive and Negative Affect Schedule (PANAS), Cognitive Reflection Task (CRT), Interpersonal Reactivity Index (IRI), Empathy Assessment Index (EAI), Psychological Entitlement Scale (PES), 18-point Need for Cognition (NFC), and a basic demographic survey were administered. Time-series clustering using the Dynamic Time Warping (DTW) algorithm was applied to determine difference is driving styles. The use of pre-driving psychometric questionnaires to determine the most suitable metrics for predicting driving style outside of the automobile suggest promising results. The results of the binary logistic regression indicate that an individual's annual mileage, crash history, NFC total score, EAI affective response, EAI empathic attitude, and education level, contribute significantly towards predicting their driving style.

Keywords: driving simulator; mental workload; time-series clustering; engagement level; psychometric questionnaires.

[1] * Corresponding author. Tel.: +001-785-312-0845;
E-mail address: kummetha@usf.edu

1. Introduction

Driver classification is a common strategy used to easily group individuals based on similar behavioral traits, driving style, or demographic characteristics. Classification provides an efficient approach to isolating unique traits associated with specific driver types under various driving conditions. This process can not only be used to develop better driver behavioral models for car-following or lane changing, but also to identify risk-averse driving behaviors; thus, providing context for targeted safe driving education.

Several studies use classification to identify behavior and driving styles [1-3]. Lin et al., classified ten drivers by analyzing physiological (*defined as the information pertaining to a driver's bodily functions such as cognition, heart rate, vision-related changes, skin conductance, sweat, and others*) measures in response to an unexpected obstacle in a virtual simulator. Drivers were classified using driving trajectory, speeds, and event-related potentials (ERP) to form two categories: aggressive and gentle [3]. Kondyli and Elefteriadou observed trends suggesting three driver behavioral types: aggressive, average, and conservative. Aggressive drivers were observed to drive at high speeds (>15 mph speed limit), perform six discretionary lane changes, and aggressive lane changing [4]. Average drivers did not exceed 10 mph over posted speed limits while conservative drivers maintained speeds within 5 mph of posted limits.

Previous studies indicate that driver classification has primarily been performed using aggregate metrics and questionnaires. Furthermore, very few studies employ both measurable physiological changes and environmental factors when assigning driving styles. To achieve a more holistic classification, this study seeks to address the above shortcomings using a data-driven approach to assess driving tasks performed under varying mental workloads (*defined as the proportion of mental capacity required by an individual to perform a task [5]*), including distracted driving. The main goals of this research are to classify drivers based on their simulated driving performance using a data-driven approach and predict their classification (i.e., driving profile) using various psychometric questionnaires. This is achieved by:

- Data fusion of psychophysiological and driving performance changes experienced by drivers when engaged in simulated tasks of varying difficulty, coupled with machine-learning techniques to provide a more accurate estimate of the ground truth for behavioral classification, and
- Employing pre-driving psychometric questionnaires to determine the most suitable subjective metrics for predicting driving style. If psychometric questionnaires have strong predictive qualities, this may save resources and time for defensive driving educators, researchers, driving rehabilitation specialists, and drivers.

Researchers have applied vast psychometric evaluation metrics in behavioral studies to assess mood and personality, cognitive engagement, and empathy and social decision-making. The most popular questionnaires/metrics used are described in Table 1.

Table 1: Psychometric Evaluators

Questionnaire/Metric	Description
Positive and Negative Affect Schedule (PANAS) [6]	Assess negative affect (i.e., distressed or nervous) and positive affect (i.e., excited or proud) on a scale of 1 to 5.
Cognitive Reflection Task (CRT) [7, 8]	Three questions designed to measure the ability to suppress an intuitive wrong answer in favor of a deliberative right answer.
Interpersonal Reactivity Index (IRI) [9]	Measure individual differences in empathy across 4 subscales (perspective taking, fantasy, empathic concern, and personal distress).
Empathy Assessment Index (EAI) [10]	50-item questionnaire with 5 sub-dimensions (i.e., affective response-AR, self-other awareness-SA, emotion regulation-ER, perspective taking-PT, and empathic attitudes-EA).
Psychological Entitlement Scale (PES) [11, 12]	Nine self-reported measures to quantify the stable and pervasive sense that one deserves more than others.
Need for Cognition (NFC) [13]	34 questions to assess the tendency to engage in and enjoy effortful cognitive endeavors.

2. Methodology

The methodology involves collecting driving performance and psychophysiological data from six tasks in a simulator setting. A fixed-base simulator, in a half cab, utilizing three forward screens to provide a 170º horizontal field of view and a rear screen was used in this research. Ninety participants ranging from 18 to 64 years of age were recruited for the study. For the simulated drive, six tasks consisting of select external conditions were created. Each task was eight kilometers long with a posted speed limit of 70 mph (112.7 km/h) and was designed to capture incremental changes to behavioral and driving performance measures. Tasks varied in traffic density, but drivers were free to maintain any gap or speed. The task sequence for each participant was completely randomized to eliminate order bias. Psychophysiological measures were continuously collected during the drive by capturing micro-level changes to vision and attention. At the end of each task, participants were required to complete the NASA-Task Load Index (NASA-TLX) and Situation Awareness Rating Technique (SART) questionnaires to provide a well-rounded subjective baseline. Tasks 1 to 5 were performed under conventional driving conditions while task 6 required participants to attempt a visual secondary task on a 10-inch Windows tablet, up to four times at random intervals no longer than 20 seconds each [8, 12].

For the continuous measurement of vision-related metrics such as pupil diameter, gaze point vectors, blinks, and fixations, a Fovio FX3 eye tracker with a data collection frequency of 60Hz was utilized. Mental workload was established from the eye tracker in one-second increments using a patented pupillometric technique called Index of Cognitive Activity (ICA) [14]. Engagement level (EL), defined as the level of engagement/arousal experienced in response to a particular task, was an important component in this research [15]. Enobio 8 portable Bluetooth-based system was used to capture participants baseline and time-correlated electroencephalogram (EEG) at 500Hz. Driver engagement level was calculated using the power spectral density (PSD) relationship $\beta/(\alpha+\theta)$ established by Pope et al. and Prinzel III et al., at PZ, CZ, P3, and P4 electrode locations, to assess alertness and engagement, mental effort, and attention investment [15, 16]. A single combined PSD value representing driver engagement level was obtained every 10-seconds of the drive. PSD combined values closer to 0 indicate lower engagement levels. Further, a Polar H10 chest strap was used to monitor drivers heart rate (HR) at 1Hz as a surrogate measure of mental workload. Data fusion was performed to uniformly synchronize all data variables to 10Hz.

3. Analysis and Results

Mental workload and situation awareness between individual tasks were computed using a series of repeated measure analysis of variances (ANOVAs) with task 1 as the baseline. The average NASA-TLX scores showed no significant differences (α = 0.05) between the mean scores of task 2 and task 1 (baseline). However, significant differences in scores were observed between the baseline and tasks 3, 4, 5, and 6. The significant increases in NASA-TLX scores suggest that the developed tasks captured various levels of mental workload. Similar trends across tasks were observed with the SART scores. No significant differences were observed between the baseline and the mean scores of the task 2 and task 3; however, task 4, task 5, and task 6 showed significant differences.

As tasks were constant between participants, time-series clustering using the Dynamic Time Warping (DTW) algorithm was selected for the analysis. The DTW algorithm provides a robust method of measuring and matching similar time-series datasets even if lag in time points is present [17, 18]. In this study, acceleration, steering wheel angle in degrees, headway, Mental workload as ICA from left eye, and engagement level, were used for time-series clustering. The two established clusters did not exhibit qualities of aggressive driving as evident from the average speeds and headways, with average speeds within 6 km/h of the posted speed limit and the smallest headway at 68.4m (2.2 seconds). Out of 83 drivers that completed all the components of the questionnaires and driving tasks, 19 drivers were classified as overall conservative while 64 drivers were classified as moderate.

Binary logistic regression (0 indicating conservative and 1 indicating moderate driving style) was then performed to predict the driving style using the following survey variables: PANAS (subscales: positive affect, negative affect), CRT, IRI (subscales: perspective taking, fantasy, empathic concern, and personal distress), EAI (subscales: AR, SA, ER, PT, EA), PES, 18-point NFC, and other demographics (i.e., age; gender; annual mileage traveled; crash history; insurance type; traffic violation history; cell phone use while driving; and education level). The results from the binary logistic regression indicate that an individual's annual mileage, crash history (0-1, 2-3, >3 crashes-last 5 years), NFC total score, EAI affective response (AR), EAI empathic attitude (EA), and education level (5 factors: 1: high school; 2: current college student; 3: finished college; 4: current graduate student; 5: finished graduate school with at least a master's degree), can significantly help in predicting their driving style (moderate or conservative) as showin in Table 2.

The odds ratios indicate that an increase in annual mileage indicates a less likely chance of being a conservative driver. Having a history of crashes reduces the likelihood of being a moderate driver by just under 97%. This could be either due to conservative drivers being slow and more susceptible to higher mental workloads or the relatively small sample size used in this research. Further, an increase in the NFC score by one point can increase classification as a moderate driver by 85%. This could be because high scorers in the NFC are individuals who enjoy thinking and apply their thinking skills easily, suggesting quicker behavioral adaptation to changes resulting from changing driving conditions. EAI AR which is a subcomponent of empathy derived from affective response was also found to be significant. An increase in the EAI AR score suggests up to 67% increased likelihood of being classified as moderate driver. Although EAI EA and education level were not significant at $\alpha = 0.05$ level, they are included in the final results as excluding them substantially effects prediction accuracy.

Table 2: Logistic Regression Summary

Variable	B	S.E.	Wald	df	Sig.	Exp(B)
Annual mileage	.000	.000	4.367	1	.037*	1.000
Crash history	-3.412	1.209	7.967	1	.005**	.033
NFC total	-.095	.035	7.304	1	.007**	.909
EAI AR	-.209	.099	4.449	1	.035*	.812
EAI EA	.212	.114	3.447	1	.063	1.236
Education level	-.467	.265	3.117	1	.077	.627
Constant	9.985	2.994	11.125	1	.001**	21698.909

*<0.05, **<0.01; Cox & Snell $R^2 = 0.218$; Nagelkerke $R^2 = 0.331$

4. Conclusions

This study aimed at addressing the shortcomings in driver classification research using a data-driven approach and assess driving tasks performed under varying mental workloads. Ninety drivers were recruited to participate in both the subjective and driving components of this research. The main goals of this research were achieved by successfully performing data fusion of psychophysiological and driving performance. Time-series clustering using the DTW algorithm was applied to provide a more accurate estimate of the ground truth for behavioral classification. Further evaluating the use of pre-driving psychometric questionnaires to determine the most suitable metrics for predicting driving style outside of the automobile suggest promising results. Binary logistic regression summary indicates that an individual's annual driving mileage, crash history, NFC total score, EAI affective response, EAI empathic attitude, and education level, can be directly attributed to identifying their driving style.

Although the research goals were achieved, the limitations of simulator-based driving data should be acknowledged. However, the use of complex physiological equipment such as EEG and presence of hazardous conditions (in vehicle distractions) might not be feasible for real-world testing. A larger sample size would provide more concrete evidence of the observed findings.

Acknowledgment

The authors would like to thank the Mid-America Transportation Center (MATC) for funding [grant number 69A3551747107] and supporting this research initiative. They would also like to thank Dr. Evangelia Chrysikou from Drexel University for aiding with the questionnaires and methodological developments.

References

1. Feng, F., et al., *Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data.* Accident Analysis & Prevention, 2017. **104**: p. 125-136.
2. Murphey, Y.L., R. Milton, and L. Kiliaris. *Driver's style classification using jerk analysis*. in *2009 IEEE Workshop on Computational Intelligence in Vehicles and Vehicular Systems*. 2009.
3. Lin, C., et al. *Driving Style Classification by Analyzing EEG Responses to Unexpected Obstacle Dodging Tasks*. in *2006 IEEE International Conference on Systems, Man and Cybernetics*. 2006.
4. Kondyli, A. and L. Elefteriadou, *Modeling Driver Behavior at Freeway–Ramp Merges.* Transportation Research Record, 2011. **2249**(1): p. 29-37.
5. Brookhuis, K.A., G. de Vries, and D. de Waard, *The effects of mobile telephoning on driving performance.* Accid Anal Prev, 1991. **23**(4): p. 309-16.

6. Watson, D., L.A. Clark, and A. Tellegen, *Development and validation of brief measures of positive and negative affect: the PANAS scales.* Journal of personality and social psychology, 1988. **54**(6): p. 1063.

7. Frederick, S., *Cognitive Reflection and Decision Making.* Journal of Economic Perspectives, 2005. **19**(4): p. 25-42.

8. Kummetha, V.C., et al., *Safety analysis of work zone complexity with respect to driver characteristics — A simulator study employing performance and gaze measures.* Accident Analysis & Prevention, 2020. **142**: p. 105566.

9. Davis, M.H., *Measuring individual differences in empathy: Evidence for a multidimensional approach.* Journal of Personality and Social Psychology, 1983. **44**(1): p. 113-126.

10. Lietz, C.A., et al., *The Empathy Assessment Index (EAI): A confirmatory factor analysis of a multidimensional model of empathy.* Journal of the Society for Social Work and Research, 2011. **2**(2): p. 104-124.

11. Campbell, W.K., et al., *Psychological entitlement: interpersonal consequences and validation of a self-report measure.* J Pers Assess, 2004. **83**(1): p. 29-45.

12. Kummetha, V.C., *Incorporating Biobehavioral Architecture into Car-Following Models: A Driving Simulator Study*, in *Civil, Environmental, and Architectural Engineering*. 2020, University of Kansas.

13. Cacioppo, J.T., R.E. Petty, and C. Feng Kao, *The Efficient Assessment of Need for Cognition.* Journal of Personality Assessment, 1984. **48**(3): p. 306-307.

14. Vogels, J., V. Demberg, and J. Kray, *The Index of Cognitive Activity as a Measure of Cognitive Processing Load in Dual Task Settings.* Front Psychol, 2018. **9**: p. 2276.

15. Pope, A.T., E.H. Bogart, and D.S. Bartolome, *Biocybernetic system evaluates indices of operator engagement in automated task.* Biological Psychology, 1995. **40**(1): p. 187-195.

16. Prinzel, L.J., *Empirical Analysis of EEG and ERPs for Psychophysiological Adaptive Task Allocation*. 2001.

17. Mueen, A. and E. Keogh, *Extracting Optimal Performance from Dynamic Time Warping*, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016, Association for Computing Machinery: San Francisco, California, USA. p. 2129–2130.

18. Giorgino, T., *Computing and Visualizing Dynamic Time Warping Alignments in R: The dtw Package.* Journal of Statistical Software, 2009. **31**(7): p. 1 - 24.