U.S. Department of Transportation
National Highway Traffic Safety Administration

NHTSA

DOT HS 813 181 November 2021

Older Driver Performance Across Six Naturalistic Studies

DISCLAIMER

Suggested APA Format Citation:

Staplin, L., Lococo, K. H., Mastromatto, T., & Sifrit, K. J. (2021, November). *Older driver performance across six naturalistic studies* (Report No. DOT HS 813 181). National Highway Traffic Safety Administration.

Technical Report Documentation Page

1. Report No. DOT HS 813 181	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Older Driver Performance Across Six Naturalistic Studies		5. Report Date November 2021 6. Performing Organization Code
7. Authors Loren Staplin, Kathy H. Lococo, Tia Mastromatto, and Kathy J. Sifrit		8. Performing Organization Report No.
9. Performing Organization Name and Address TransAnalytics, LLC 336 West Broad Street Quakertown, PA 18951		10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTNH2216D00016/0004
12. Sponsoring Agency Name and Address National Highway Traffic Safety Administration 1200 New Jersey Avenue SE Washington, DC 20590		13. Type of Report and Period Covered Final Technical Report 14. Sponsoring Agency Code NPD-320

15. Supplementary Notes
Dr. Kathy Sifrit was the NHTSA task order manager on this project.

16. Abstract
This study aggregated and analyzed naturalistic data collected from 232 participants from six NHTSA-sponsored studies to determine the extent older drivers' scores on particular clinical measures were associated with their ability to control vehicles, including responding appropriately to traffic control devices and maintaining proper lane position, as documented in on-road performance evaluations by a certified driver rehabilitation specialist. Researchers also examined the association between the clinical measures and indices of driving exposure, including total mileage and total driving time, and exposure to high-speed and limited-access roadways and other challenging situations. Analyses of NHTSA on-road performance and naturalistic driving data explored whether participants with poorer driving skills were more likely to limit their overall driving (time and/or miles), avoid potentially difficult conditions, or otherwise self-regulate appropriately. In addition to the combined NHTSA data analyses, researchers analyzed data from the SHRP2 Naturalistic Driving Study for a sample of 1,045 drivers 60 and older, including driver performance and exposure measures, and measures of driver (cognitive and physical) functional status. Results from the NHTSA and SHRP2 data analyses indicated only "very weak" to "weak" relationships between the predictor and criterion variables. Exploratory analyses of the SHRP2 data yielded additional findings suggesting significant differences in drivers' responses during crash and near-crash events tied to measures of both cognitive and physical function.

17. Key Words older driver, functional ability, driving performance, driving exposure, safety, SHRP2, naturalistic		18. Distribution Statement Document is available to the public from the National Technical Information Service, www.ntis.gov.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 42	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Table of Contents

Executive Summary

Over the previous decade NHTSA has sponsored a series of investigations aimed at helping normally aging older drivers stay safe behind the wheel. Data collected for these studies included measures of the functional (especially cognitive) status of older drivers, their performance of critical strategic, tactical, and operational driving tasks, and their driving habits, including how much they drive. These studies had relatively small sample sizes—20 to 60 participants—with driving exposure (how much, and under what conditions participants drove) data collection lasting up to a month. As a group these studies shared some important methodological characteristics. There was considerable overlap in measures of driver functional status, on-road performance, and in instrumented vehicle data collected across these studies. Together, these attributes provided a potential for data synthesis, as combining data sets from the individual studies could provide greater statistical power for testing relationships among driver functional status, driving performance, and driving exposure.

Early in the same timeframe, the Naturalistic Driving Study (NDS) undertaken as part of the second Strategic Highway Research Program (SHRP2) enrolled 1,045 drivers 60 and older, and their cars were instrumented for naturalistic data collection that lasted 1 to 2 years. In addition, the SHRP2 NDS participants underwent functional screening using the same computer-based program applied in the NHTSA studies. By analyzing selected SHRP2 data researchers may gain further insight into the same relationships that formed the premise for the NHTSA data synthesis. In addition, comparisons of the combined NHTSA and the SHRP2 data sources could serve to gauge the reliability of data collected in smaller, cross-sectional investigations.

The NHTSA data synthesis aggregated performance and exposure variables based on their availability across studies. After applying screening criteria to the 52 possible measures, a single metric of driver performance (the *total road test error score* on an on-road driver evaluation), 30 driving exposure measures, and three functional status predictors (two cognitive, one physical) met the criteria to serve as the bases for aggregating data sets acquired in six different NHTSA projects.

Researchers specified data for extraction from the NDS to overlap to as great an extent as possible with the measures qualified for analysis in the combined data synthesis. This effort was successful with respect to the functional predictors and measures of driving exposure, but none of the kinematic-based performance measures in the NDS offered a valid comparison with the CDRS's context-based driver error score. The NDS data set, however, provided direct safety measures (crashes and near-crashes) not available in the combined data set, to be used in the planned analyses as outcome/criterion variables.

Correlations between functional measures and road test scores, road test scores and driving exposure measures, and functional and exposure measures in the aggregated NHTSA data were "very weak" (.00 - .19), with a few exceptions that were "weak" (.20 - .39). At best, researchers could account for only 10% of the variance in these analyses, and that was for the association between scores on one of the functional (cognitive) status measures and driver age, so was not associated with any exposure or performance measure. However, among the isolated weak correlations involving functional status measures, road test error scores, and exposure, the *direction* of the correlation associated both lower maximum average trip speed and worse behind-the-wheel performance (higher error scores) with poorer functional test scores.

The SHRP2 analysis results were more informative in several respects. First, the persistent weak relationships between functional status and exposure reinforced the outcome of the aggregated NHTSA data analyses, which were based on far fewer drivers. And for both data sets, the interpretation of results is consistent with the conclusion expressed in previous NHTSA research that older adults' choices about when, where, and how often they drive primarily reflect habits acquired from many years of experience keyed to accessing varied goods and services, and maintaining social connections within the community, rather than being grounded in an individual's current level of fitness.

The strongest contribution of the SHRP2-based data was in exploratory analyses of how drivers' vehicle control responses in crash and near-crash situations were related to differences in particular functional abilities. Specifically, the time from the instant a driver could perceive a "precipitating event" to the start of the driver's reaction with a vehicle control response was significantly longer (slower) for participants whose functional test scores indicated serious impairment on one or more measures keyed to visual information processing speed and visual search with divided attention—but only when a crash or near-crash was related to a conflict that originated in the driver's *peripheral* visual field (e.g., a car approaching on an intersecting roadway). Research has shown that older drivers scan left and right less frequently than other drivers, which could make them more vulnerable to conflicts arising in their periphery (Romoser et al., 2013; Dukic & Broberg, 2012; Bao & Boyle, 2009; and Scott et al., 2013).

Another significant finding in the SHRP2-based analyses was that the time between the driver's reaction and the time of impact was, on average, significantly shorter for drivers *with* serious impairment in lower limb mobility than for those *without* serious impairment. Such an impairment could be expected to result in further travel toward the point of impact before the driver fully executes the control movement (e.g., braking).

Overall, the SHRP2 analysis results are consistent with some revealed in the combined data, but extend beyond those findings to affirm researchers' models of how age-related changes in function (particularly cognition) may affect driving performance. A discussion of the study's findings highlights the need to frame research questions precisely in future studies that seek to understand how the capabilities and limitations of normally aging adults may influence the crash avoidance behaviors of these drivers.

Introduction

Background

Conventional wisdom holds that older drivers reduce their crash risk by self-regulating, limiting their driving to situations they believe they can navigate safely. However, crash data show that drivers 70 and older are more likely than middle-aged drivers to contribute to crashes, particularly at intersections, turning left across traffic, and when driving at highway speeds (Stutts et. al, 2009). The risk increases for drivers 80 and older. In the Stutts study the risk of involvement in fatal two-vehicle crashes for drivers 70 to 79 was nearly equivalent to that of teenage drivers, and for those 80 and older it was four times higher than expected based on (induced) driving exposure.

In 2018 older drivers made up 20% of all licensed drivers but only 14% of drivers involved in fatal crashes (National Center for Statistics and Analysis, 2020). This suggests that, despite functional declines associated with aging, older drivers appear to generally adapt to those changes, and/or use strategies to reduce their risk so as to be under-represented in fatal crashes. Crash rates based on population are likely to understate the risk per mile traveled for those who drive the fewest miles (generally the oldest and youngest drivers) and overstate it for those who drive the most. The fatal crash involvement rates within the population of drivers 65 and older were highest for those 85 and older.

A better understanding of the relationship between the functional (especially cognitive) status of older drivers, their performance of critical strategic, tactical, and operational driving tasks, and the extent to which they self-regulate their exposure to risky driving situations is key to informing policy and developing countermeasures to help these motorists keep driving safely. Since 2010 NHTSA has sponsored studies in which measures for each of the variables were obtained to address research questions related to compliance with license restrictions, the effects of medical conditions including mild cognitive impairment, the effects of physical fitness on driving performance and exposure, and driver self-regulation. Data collected from these previous studies supported analyses in the current study of how differences in functional status are related to differences in driver performance and/or exposure, and whether deficiencies in drivers' skills are associated with more restrictive driving habits.

Each of the earlier studies applied common, objective, and standardized functional screening procedures, "gold-standard" on-road performance evaluations by certified driver rehabilitation specialist (CDRS), and naturalistic measures of driving exposure. However, study limitations included: (1) geography, as data were collected only in Virginia and the Carolinas; (2) small sample sizes, typically 20 to 60 older drivers; and (3) naturalistic driving observations typically lasting only 2 to 4 weeks.

A contemporaneous research initiative, the NDS undertaken as part of the SHRP2, enrolled drivers from six States, distributed across the country, whose own cars were instrumented with cameras, radar, GPS, and other sensors. Among the study sample were 1,045 drivers 60 and older, with a period of study participation averaging 449 days. This cohort was involved in over 1,100 crash or near-crash events. Multiple measures of driving exposure were recorded or could be derived from the instrumented vehicle data, including those obtained in the NHTSA-sponsored research. In addition, the SHRP2 participants underwent functional screening using the same computer-based program for selected measures common to the NHTSA studies, including the Trail-Making Test (TMT) and the Rapid Pace Walk Test.

The results of these investigations present an opportunity for NHTSA to reexamine questions about whether older drivers with clinically significant losses in abilities previously validated as predictors of crash risk demonstrate poorer performance behind the wheel, and/or alter their driving habits appropriately, benefitting from increased statistical power for such analyses where it is possible to aggregate data across studies. At the same time, there is a rare opportunity to check on how reliable and representative the small-sample data are, using as a yardstick the much more extensive and comprehensive (albeit orders of magnitude more expensive and time-consuming to obtain) results yielded by SHRP2. This comparison could provide important information to NHTSA for defining the scope of future driver performance data collection efforts.

Objectives

This study had two objectives, to combine and analyze data from the NHTSA studies (the combined data set) to obtain more power to detect relationships among functional abilities, driving performance and habits, and exposure, and to analyze SHRP2 data to identify relationships among age, clinical scores, and driving exposure, then to compare the findings from the NHTSA and SHRP2 data. Specific research questions of interest in this study follow.

Which clinical measures predict a participant's ability to control the vehicle, including the ability to respond appropriately to traffic control signals, to other drivers, to use turn signals appropriately, or to stop in a proper position at an intersection?

What is the relationship between specific clinical measures and participants' total mileage, total driving time, and trip characteristics?

Do clinical measures predict participants' likelihood to drive on high-speed and/or limited access roadways, or the time of day that participants drive; e.g., the probability they will drive at night or during rush hour?

Do relationships between scores from the driver evaluation and naturalistic data suggest that drivers with poorer driving skills are more likely to limit their overall driving (time and/or miles), avoid potentially difficult conditions such as night or rush hour driving, or otherwise self-regulate appropriately?

Do the findings from the NHTSA-collected data (shorter naturalistic data collection interval) differ from those based on the SHRP2 data (longer naturalistic data collection interval)? Were there any limitations associated with the NHTSA data being primarily collected from drivers in just three States and over a relatively short data collection period?

A better understanding of the extent to which older drivers match their functional abilities to their choices of when and where to drive may benefit clinicians who work with older adults. A better understanding of the relationship between scores on formal driver evaluations and driving exposure should enhance driver evaluators' ability to provide recommendations to improve the safety of their clients. Such information may also inform educational material for older drivers and their families. Analyses of the SHRP2 data may further NHTSA's understanding of the relationship between functional measures, driving habits, and exposure based on a much longer exposure period, possibly revealing relationships that the combined data alone did not.

Methods

Aggregating NHTSA Study Data

The researchers aggregated data across six NHTSA-sponsored studies, to the extent permitted by data availability and with consideration of the similarity of the data sets in terms of the range and distribution of values for a given measure. The citations for these six studies are listed below.

Joyce, J., Lococo, K. H., Gish, K. W., Mastromatto, T., Stutts, J., Thomas, D., & Blomberg, R. (2018, April). *Older driver compliance with license restrictions* (Report No. DOT HS 812 486). National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/36716/dot_36716_DS1.pdf?

Staplin, L., Lococo, K. H., Mastromatto, T., Gish, K. W., Golembiewski, G., & Sifrit, K. J. (2019a, January). *Mild cognitive impairment and driving performance* (Report No. DOT HS 812 577). National Highway Traffic Safety Administration. www.nhtsa.gov/sites/nhtsa.gov/files/documents/13687-cognitive_impairment_report_010419_v4-tag.pdf

Staplin, L., Mastromatto, T., Lococo, K. H., Gish, K. W., & Brooks, J. O. (2017, August). *The effects of medical conditions on driving performance* (Report No. DOT HS 812 439). National Highway Traffic Safety Administration. www.nhtsa.gov/sites/nhtsa.gov/files/documents/812439_the_effects_medical_conditions_driving_performance_0.pdf

Thomas, F. D., Graham, L. A., Finstad, K. A., Wright, T. J., Blomberg, R. D., Lococo, K., Gish, K., Staplin, L., Stutts, J., Wilkins, J., Crompton, C., & Sifrit, K. J. (2020, April). *Older drivers' self-regulation and exposure* (Report No. DOT HS 812 930). National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/49163/dot_49163_DS1.pdf

Staplin, L., Lococo, K. H., Gish, K. W., Stutts, J., & Srinivasan, R. (2019b, September). *Activity level, performance and exposure among older drivers* (Report No. DOT HS 812 734). Washington, DC: National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/42285/dot_42285_DS1.pdf

Staplin, L., Lococo, K. H., Stutts, J., Robison, K., Levitt, D., Srinivasan, R., & Sifrit, K. J. (2021,May). *Physical fitness training and older driver performance and exposure* (Report No. DOT HS 813 107). National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/55934/dot_55934_DS1.pdf

Candidate variables for data aggregation across studies included *clinical/functional ability* measures, measures of *driver performance*, and *driving exposure* (how much and under what conditions participants drove) measures. Each of these domains included specific data elements, some of which were included in all six data sets while others were not. The researchers considered for aggregation only the variables for which data were collected for *more than 50% of the pooled participants*.

Table 1 presents variable descriptions, sample sizes, and data availability across the NHTSA studies for each of the variables considered for inclusion. In the aggregate, the studies included a total of 232 participants. Demographic information (age, sex) was available for all participants (data on participants' race and ethnicity were not collected), as were all the time-based exposure measures (e.g., trip counts, duration, time-of-day, trips during adverse weather). Among the functional ability measures, data describing cognitive status (TMT and the Snellgrove Maze Test) were available for most participants as was a measure of physical status (lower limb strength/flexibility). However, measures of vision and visual attention (acuity, contrast sensitivity, useful field of view) and simple brake reaction time failed to meet the inclusion criterion.

Within the driver exposure domain, there was mixed availability of measures of speed and distance traveled. These measures were derived from GPS data, which were more often missing and/or of questionable reliability than were the exposure data elements derived through video coding. Thus, the measures describing percentages of trips with a given radius of miles from home and the percentage of miles driven on limited access highways were unavailable for most participants. Finally, data indicating the presence of passengers, while not collected in most of the *studies*, were available for most *participants*, as these data were available for participants in the studies with the largest sample sizes.

Table 1: Variables Available for Analysis by NHTSA Study

Variable Names and Description	Variables Available for Analysis					
	Joyce et al. (2018)	Staplinet al. (2019b)	Staplin et al. (2017)	Thomas et al. (2020)	Staplin et al. (2019a)	Staplin et al. (2021)
SN: Subject number specific to original NHTSA study	✓	✓	✓	✓	✓	✓
Age: Subject age at the time data were collected	✓	✓	✓	✓	✓	✓
Sex: Subject sex: M= Male, F=Female	✓	✓	✓	✓	✓	✓
OP_PointsOff: Number of points scored off for Operational Skills		✓		✓	✓	✓
TA_PointsOff: Number of points scored off for Tactical Skills		✓		✓	✓	✓
ST_PointsOff: Number of points scored off for Strategic Skills		✓		✓	✓	✓
BTW_TotalPointsOff: Total number of points scored off (OP + TA + ST)		✓		✓	✓	✓
RoadTestGrade: A (0-24 points off, pass with no restrictions), B (25-49 points off, pass with recommendations), C (50-75 points off, marginal with restrictions), D (76-99 points off, fail) or F (100+ points off, fail)		✓		✓	✓	✓
RapidPaceWalk: Time (seconds): Participant walked 10 feet, turned, and walked back 10 feet for a total distance of 20 feet.		✓	✓		✓	✓
Trails_A: Time to complete Trail Making Test A (seconds)		✓	✓	✓	✓	✓
Trails_B: Time to complete Trail Making Test B (seconds)		✓	✓	✓	✓	✓
Maze_1: Time to complete Maze 1 (seconds)		✓		✓	✓	✓
Maze_2: Time to complete Maze 2 (seconds)		✓		✓	✓	✓
Maze_Both: Total Time to complete Mazes 1 and 2 (seconds)		✓		✓	✓	✓
UFOV2: Useful Field of View Subtest 2, Divided Attention (100 to 333 milliseconds)			✓	✓		
CS: Contrast Sensitivity Score (Log of CS)		✓	✓	✓		
Acuity_Far: Far Acuity Score (Log Mar Score)		✓	✓	✓		
Acuity_Near: Near Acuity Score (Log Mar Score)			✓	✓		
Simple_Brake_RT: Reaction Time (seconds)		✓*	✓	✓		
TripCount: Number of trips taken by subject	✓	✓	✓	✓	✓	✓
VehEquipDays: The number of days the subject's vehicle was equipped with video and GPS equipment	✓	✓	✓	✓	✓	✓
ShortestTripTime_Mins: Shortest trip time (minutes)	✓	✓	✓	✓	✓	✓
LongestTripTime_Mins: Longest trip time (minutes)	✓	✓	✓	✓	✓	✓

Variable Names and Description	Variables Available for Analysis					
	Joyce et al. (2018)	Staplinet al. (2019b)	Staplin et al. (2017)	Thomas et al. (2020)	Staplin et al. (2019a)	Staplin et al. (2021)
AverageTripTime_Mins: Mean of all trip times (minutes)	✓	✓	✓	✓	✓	✓
TotalDrivingTime_Mins: Total time the subject drove during the exposure study (minutes)	✓	✓	✓	✓	✓	✓
ShortestTripDist_Miles: Shortest trip distance (miles)	✓	✓	✓	✓	✓	✓
LongestTripDist_Miles: Longest trip distance (Miles)	✓	✓	✓	✓	✓	✓
AverageTripDist_Miles: Mean trip distance (Miles)	✓	✓	✓	✓	✓	✓
TotalDrivingDistance_Miles: Total driving distance during the exposure study (Miles)	✓	✓	✓	✓	✓	✓
TripCountPerDay: Number of trips per vehicle instrumentation day (TripCount/VehEquipDays)	✓	✓	✓	✓	✓	✓
MinutesPerDay: Number of minutes the subject drove per vehicle instrumentation day (TotalDrivingTime_Mins/VehEquipDays)	✓	✓	✓	✓	✓	✓
MilesPerDay: Number of miles per vehicle instrumentation day (TotalDrivingDist_Miles/VehEqipDays)	✓	✓	✓	✓	✓	✓
MaxSpeed: The highest speed reached during all trips (mph)	✓	✓		✓	✓	✓
PercentRadiusHome_1: % of Trips ≤ 1 Mile Radius of Home	✓	✓				
PercentRadiusHome_1-5: % of Trips >1 and ≤ 5 Miles Radius of Home	✓	✓				
PercentRadiusHome_5-20: % of trips >5 and ≤ 20 Miles Radius of Home	✓	✓				
PercentRadiusHome_20: % of trips >20 Miles Radius of Home	✓	✓				
PercentMilesLimitedAccess: % of miles driven on limited access roadways or roadways with 55+ mph speed limits. Joyce et al. (2018) and Staplin et al. (2017) def: speeds 55+, concrete dividers, ramps for on/off access, no intersections. Staplin et al. (2019b) def: roads with speeds of 60+ mph speed limits.	✓	✓	✓			
PercentTrips_6AM-10AM: % of trips started ≥ 6 a.m. and < 10 a.m.	✓	✓	✓	✓	✓	✓
PercentTrips_10AM-3PM: % of trips started ≥ 10 a.m. and < 3 p.m.	✓	✓	✓	✓	✓	✓
PercentTrips_3PM-8PM: % of trips started ≥ 3 p.m. and < 8 p.m.	✓	✓	✓	✓	✓	✓
PercentTrips_8PM-6AM: % of trips started ≥ 8 p.m. and < 6 a.m.	✓	✓	✓	✓	✓	✓
PercentTrips_Under1mi: % of trips ≤ 1 mile total distance	✓	✓	✓	✓	✓	✓
PercentTrips_1-2.5mi: % of trips > 1 mile and ≤ 2.5 miles total distance	✓	✓	✓	✓	✓	✓
PercentTrips_2.5-5mi: % of trips > 2.5 miles and ≤ 5 miles total distance	✓	✓	✓	✓	✓	✓

Variable Names and Description	Variables Available for Analysis					
	Joyce et al. (2018)	Staplinet al. (2019b)	Staplin et al. (2017)	Thomas et al. (2020)	Staplin et al. (2019a)	Staplin et al. (2021)
PercentTrips_5-10mi: % of trips > 5 miles and ≤ 10 miles total distance	✓	✓	✓	✓	✓	✓
PercentTrips_10-20mi: % of trips > 10 miles and ≤ 20 miles total distance	✓	✓	✓	✓	✓	✓
PercentTrips_Over20mi: % of trips longer than 20 miles total distance	✓	✓	✓	✓	✓	✓
PercentTrips_Night: % of trips made at night. For Joyce et al. (2018), defined as trips coded as dusk, dark, and dawn. For Staplin et al. (2017), defined as trips from 8 pm to 6 am. For Staplin et al. (2019a), Thomas et al. (2020), Staplin et al. (2019b), and Staplin et al. (in press), the video coder identified trips made at nighttime as those with dark skies at dusk through dawn (i.e., low contrast conditions)	✓	✓	✓	✓	✓	✓
PercentTrips_Rain_Fog_Wet: % of trips made in adverse weather, including rain, fog, snow, sleet, or on wet pavement	✓	✓	✓	✓	✓	✓
PercentTrips_Passengers: % of trips with 1 or more passengers	✓			✓	✓	
PercentTrips_RushHour: % of trips begun at rush hour (defined as 6-9 am and 4-7 pm)	✓	✓	✓	✓	✓	✓
TripSpeedAvg_Min: The lowest mean speed in MPH from all trips	✓	✓		✓	✓	✓
TripSpeedAvg_Max: The highest mean speed in MPH from all trips	✓	✓		✓	✓	✓
TripSpeedAvg_Avg: The mean of the mean speed in MPH from all trips	✓	✓		✓	✓	✓

Note: * For 17 of 29 participants

After identifying candidates for data aggregation based on their availability across studies, the researchers examined the similarity of data planned for inclusion as independent/ predictor variables in these analyses with respect to range, variability, and measures of central tendency (for continuous measures). Clinical measures analyzed in relation to both performance and exposure included the Rapid Pace Walk, TMT Part A and Part B, and the Snellgrove Maze Test (Mazes 1 and 2 individually, and their combined score). Performance measures analyzed in relation to exposure included behind-the-wheel road test scores for tactical points off, strategic points off, and total points off.

Appendix A shows the data distributions for these measures using summary tables and plots. There were clear differences in mean and median performance levels across studies, particularly for the measures of cognitive function. Sometimes, though not always, performance for Staplin et al. (2019b) was most discrepant from the others, which is to be expected given that the study required recruiting participants with relatively poor scores on cognitive measures. Still, the data distributions for the Rapid Pace Walk, TMT, and Maze Test measures manifested similar (though not congruent) profiles across the data collection efforts described above. That is, the probability of a given score rose and fell synchronously across each of the studies. This characteristic provided support for retaining all independent variables in the analyses of predictor-criterion relationships for the aggregated NHTSA data, hereafter referred to as the combined data set.

The summary tables and plots relating to Road Test Scores were less convincing. Road Test Scores were a performance measure planned as a criterion variable in relation to driver functional ability, and as a predictor variable in relation to driver exposure. Again, the data for Staplin et al. (2019b) were most discrepant, but much more so than with respect to the functional measures. The indices of central tendency and variability are markedly *lower* for this measure than for the other functional measures for *tactical* points off, and much *higher* for *strategic* points off. Two possible explanations are *1)* the sample was selected to include drivers with mild cognitive impairment (MCI), and thus differed from the other studies by design; and *2)* the CDRS evaluators who scored performance for Staplin et al. (2019b) were different than for the three other studies and, while working from the same (Modified Miller) scoring template and trained by the CDRS who participated in previous studies, may have judged an observed behavior differently. Given the inconsistencies in scores for the Road Test sub-scales in the aggregation of data collected in the NHTSA-sponsored studies, the researchers opted to rely on *total points off* as the exclusive measure of driver performance.

In summary, the aggregation of data from the six NHTSA data collections required data availability from a majority of the (pooled) participants. The functional ability measures for which data were aggregated were the *Rapid Pace Walk*, *TMT Part A* and *Part B*, and the *Snellgrove Maze Test* (Mazes 1 and 2, individually, and/or their combined score). The driver performance measure for which data were aggregated was *total points off* on the CDRS-administered road test. Within this framework, researchers aggregated data for all qualifying variables checked in Table 1.

SHRP2 Variable Selection and Data Extraction

The researchers prepared and submitted a data use license (DUL) application to the SHRP2 data custodian requesting extraction of specific variables from the SHRP2 data set, to support analyses comparing relationships within this data set with those revealed in the analyses of the

NHTSA data set. The researchers submitted the study protocol for review and approval by an approved IRB, who determined that the research project did not require IRB oversight.

Among the functional ability measures obtained for drivers enrolled in the SHRP2 study, only the *Rapid Pace Walk* and *TMT Part A* and *Part B* were available as independent (predictor) variables for the planned comparisons. While many additional, potentially interesting measures of drivers' functional status were obtained in the SHRP2 study (e.g., useful field of view, contrast sensitivity), they were not obtained in the NHTSA-sponsored research for most participants.

In contrast, the SHRP2 variables available for analyses of the relationship between functional ability and exposure overlapped extensively with the exposure variables included in the NHTSA data set. Some variables were extracted directly from a SHRP2 data table (e.g., Trip Summary Table, Physical Strength Tests, Visual and Cognitive Tests), while others were derived through calculations performed by the data custodian and/or the researchers. For *percentage of trips at night*, the researchers specified the same operational definition of "night" used in Staplin et al. (2017) for the SHRP2 data extraction: trips started after 8 p.m. and before 6 a.m. This is because the only time a nighttime visibility condition was directly coded into the SHRP2 data set was when an event (either crash or baseline) occurred, which led to a video review by the data custodian's analyst.

A few omissions in the SHRP2 exposure data also deserve mention. Adverse/wet weather conditions, like nighttime driving, was only directly coded for the handful (if any) trips for a given driver where an event occurred; attempts to use a surrogate, "windshield wiper activation," in a previous SHRP2 analyses revealed that this variable was not reliable. Two other variables of interest—the percentage of trips within radius of (a given number of) miles from home and the percentage of trips on roads with posted speed greater than (a given speed), or, alternately, percentage of trips on limited access highways—were not feasible to extract for these analyses. The specific locational information needed as a reference to derive these measures resides in a data set separate from the SHRP2 data, and for this project, it was not feasible to derive these measures.

The researchers were not able to link any of the vehicle kinematic measures in the SHRP2 data set to the road test result—*total points off*—used as the metric of driver performance in the combined data set. Nor could any of the multi-level Event Severity data in the SHRP2 data set be compared to variables available in the combined data set. With respect to the SHRP2 kinematic measures, indicators of a driver's steering, braking, or acceleration behavior captured by the vehicle sensors could connote behaviors that a CDRS in the NHTSA studies might have interpreted as a (tactical) driving error, or as an appropriate response to a traffic event. In these cases, the CDRS' determination of "points off" scoring on a road test would be dependent on the context in which the behavior occurred. Barring a review of the video recorded at the individual trip level, no such contextual information was available for the kinematic data.

The researchers also explored the possibility of comparing rates of "near-crashes" and/or "crash-relevant" behaviors for older drivers within the SHRP2 data set with the rate of incidences in the NHTSA studies where the CDRS was required to apply the brake or otherwise intervene during an on-road test to avoid a conflict or (potentially) a crash. However, the data revealed only a single time that the CDRS intervened across all the participants in all the NHTSA data collection efforts.

The researchers requested that the same filters applied to all but one of the NHTSA naturalistic data sets be applied during SHRP2 data extraction:

- filter out trips where max speed is < 3 mph (apply to the raw GPS data);
- filter out trips where total distance is < 0.1 miles; and
- filter out trips where total duration is < 1 minute.

In summary, the predictor-criterion relationships using SHRP2 data to analyze for comparison to the results obtained from analyses of the combined data set were limited to the associations between age, the functional ability measures Rapid Pace Walk, and TMT Part A and Part B, and 29 of the driver exposure measures previously indicated in Table 1.

Results

The analysis outcomes reported below initially focus on the combined data set, and then examine relationships between functional ability predictors and performance and exposure outcome measures captured in the SHRP2 data for drivers 60 and older. In the latter set of analyses, researchers proceeded in a step-wise fashion, culminating with a selective narrowing of outcome event (crash and near-crash) categories to better match measurement constructs embodied in the functional predictor variables.

Aggregated NHTSA Study Sample

The combined data set was comprised of data from 232 drivers aggregated across the six separate data collection efforts. Participants included 111 females and 121 males ranging in age from 61 to 91, who altogether made a total of 17,790 trips, driving for 3,363 hours and a total of 94,684 miles. Table 2 presents summary statistics across this sample for the variables included in the analyses.

Table 2: Summary Statistics for Variables in the NHTSA Data

Variable	N	Min	Max	Mean	SD	Median
Driver Age	232	61.0	90.7	76.1	6.6	75.6
Functional Ability Measures						
BTW Total Points Off	188	0	186	33.8	32.8	23.0
Rapid Pace Walk Time (seconds)	145	3.1	13.7	6.4	1.8	6.0
Trails A Time (seconds)	209	18.0	164.0	39.9	19.3	35.4
Trails B Time (seconds)	208	41.1	300.0	106.6	50.2	92.3
Maze 1 Time (seconds)	187	2.0	145.3	26.9	30.0	12.2
Maze 2 Time (seconds)	190	4.0	163.4	21.2	17.8	16.1
Maze 1 + Maze 2 Time	187	7.0	243.1	48.1	39.1	33.7
Driving Exposure						
Trip Count	232	6	316	76.7	54.8	66.0
Vehicle Equipment Days	232	7	62	31.6	8.3	31.0
Shortest Trip Time (minutes)	232	0.0	7.2	1.5	1.0	1.1
Longest Trip Time (minutes)	232	8.8	1113.0	59.4	79.9	45.3
Mean Trip Time (minutes)	232	3.2	34.5	12.7	4.6	12.1
Total Driving Time (minutes)	232	79.9	3164.3	869.6	624.6	601.8
Shortest Trip Distance (miles)	231	0.0	2.5	0.2	0.3	0.1
Longest Trip Distance (miles)	231	1.6	328.2	36.6	40.3	26.0
Mean Trip Distance (miles)	231	0.6	26.3	5.8	3.3	5.3
Total Driving Distance (miles)	231	11.6	1859.0	409.9	288.7	347.1
Trip Count per Day	232	0.2	8.3	2.6	1.8	2.2
Minutes Driven per Day	232	2.4	143.8	29.8	22.9	23.5
Miles Driven per Day	231	0.3	84.5	14.1	10.7	12.5
Maximum Speed (mph)	212	38.5	91.0	71.4	9.4	71.8
Percent Trips 6 a.m. – 10 a.m.	231	0.0	69.7	11.4	13.7	6.7
Percent Trips 10 a.m. – 3 p.m.	231	0.0	75.6	37.6	20.4	40.0
Percent Trips 3 p.m. – 8 p.m.	231	0.0	92.9	39.4	21.6	34.8
Percent Trips 8 p.m. – 6 a.m.	231	0.0	70.8	11.6	14.0	5.9
Percent Trips Under 1 mile	231	0.0	81.7	22.3	16.5	19.0
Percent Trips 1 – 2.5 miles	231	0.0	76.9	20.8	14.0	18.6

Variable	N	Min	Max	Mean	SD	Median
Percent Trips 2.5 – 5 miles	231	0.0	67.9	23.5	12.3	21.7
Percent Trips 5 – 10 miles	231	0.0	81.0	17.1	13.3	15.3
Percent Trips 10 – 20 miles	231	0.0	60.5	11.2	11.5	7.5
Percent Trips Over 20 miles	231	0.0	48.0	5.0	6.9	3.1
Percent Trip Night	232	0.0	29.8	4.4	6.0	1.8
Percent Trips Rain/Fog/Wet	232	0.0	39.5	6.6	6.8	4.9
Percent Trips With Passengers	153	0.0	100.0	31.3	25.7	29.6
Percent Trips Rush Hour	231	0.0	78.3	30.6	14.4	28.0
Trip Speed Average (Minimum)	212	0.1	20.8	5.6	4.6	5.6
Trip Speed Average (Maximum)	212	17.6	76.0	46.6	11.6	45.6
Trip Speed Average (Mean)	212	6.4	44.4	23.5	5.5	23.2

The analyst performed correlations when the number of participants (sample size) completing both measures was at least 50% of the total number of participants in the NHTSA data set (i.e., 50% of 232 = 116). Correlation strength is classified according to the absolute value of r: very weak (.00 - .19), weak (.20 - .39), moderate (.40 - .59), strong (.60 - .79), and very strong (.80 - 1.0) (Evans, 1996). Most correlations were very weak with a few exceptions which were weak, as highlighted in bold in Table 3.

Table 3: Correlations Between Age, Functional Ability, Road Test Scores, and Driving Exposure

Functional Ability and Driving Exposure Measures	Road Test Performance	Functional Ability Measures						Driver Age
	BTW Total Points Off	Rapid Pace Walk Time	Trails A Time	Trails B Time	Maze 1 Time	Maze 2 Time	Mazes 1 + 2 Time	
Functional Ability Measures								
Rapid Pace Walk Time	(124) 0.16	---	---	---	---	---	---	(145) 0.11
Trails A Time	(187) ***0.20***	---	---	---	---	---	---	(209) ***0.26***
Trails B Time	(187) ***0.24***	---	---	---	---	---	---	(208) ***0.26***
Maze 1 Time	(185) 0.01	---	---	---	---	---	---	(187) ***0.29***
Maze 2 Time	(187) ***0.24***	---	---	---	---	---	---	(190) ***0.24***
Maze 1 + 2 Time	(185) 0.12	---	---	---	---	---	---	(187) ***0.33***
Driving Exposure Measures								
Shortest Trip Time	(188) 0.13	(145) 0.04	(209) 0.07	(208) 0.04	(187) -0.04	(190) 0.09	(187) 0.01	(232) -0.03
Longest Trip Time	(188) -0.03	(145) -0.06	(209) -0.06	(208) 0.00	(187) -0.11	(190) -0.08	(187) -0.11	(232) -0.03
Mean Trip Time	(188) -0.10	(145) -0.07	(209) 0.00	(208) 0.00	(187) 0.03	(190) -0.12	(187) -0.03	(232) -0.10
Total Driving Time	(188) -0.02	(145) -0.06	(209) -0.04	(208) -0.02	(187) -0.08	(190) -0.00	(187) -0.06	(232) ***-0.24***
Shortest Trip Distance	(187) -0.01	(144) -0.07	(208) -0.09	(207) -0.08	(186) -0.07	(189) -0.07	(186) -0.08	(231) -0.08
Longest Trip Distance	(187) -0.04	(144) -0.09	(208) -0.10	(207) -0.05	(186) -0.12	(189) -0.12	(186) -0.14	(231) -0.09
Mean Trip Distance	(187) -0.10	(144) -0.16	(208) -0.03	(207) 0.03)	(186) 0.02	(189) -0.16	(186) -0.05	(231) -0.09
Total Driving Distance	(187) 0.14	(144) -0.11	(208) -0.02	(207) 0.04	(186) -0.11	(189) 0.00	(186) -0.08	(231) ***-0.23***
Trip Count per Day	(188) -0.04	(145) -0.02	(209) -0.08	(208) -0.04	(187) -0.09	(190) 0.02	(187) -0.06	(232) ***-0.22***
Minutes Driven per Day	(188) -0.03	(145) -0.04	(209) -0.04	(208) 0.01	(187) -0.09	(190) 0.00	(187) -0.07	(232) ***-0.29***
Miles Driven per Day	(187) 0.10	(144) -0.08	(208) -0.03	(207) 0.06	(186) -0.09	(189) 0.00	(186) -0.07	(231) ***-0.28***
Maximum Speed	(187) -0.06	(144) -0.09	(189) -0.19	(188) ***-0.24***	(186) -0.07	(189) -0.15	(186) -0.12	(212) -0.12
Percent Trips 6 a.m.-10 a.m.	(188) 0.04	(145) 0.07	(209) -0.13	(208) -0.03	(187) 0.03	(190) 0.03	(187) 0.04	(231) -0.10
Percent Trips 10 a.m.-3 p.m.	(188 0.16	(145 -0.03	(209 -0.12	(208 0.01	(187 -0.02	(190 0.04	(187 0.00	(231 -0.04
Percent Trips 3 p.m.-8 p.m.	(188 -0.06	(145 0.05	(209 0.12	(208 0.00	(187 0.06	(190 -0.01	(187 0.04	(231 0.08

Functional Ability and Driving Exposure Measures	Road Test Performance	Functional Ability Measures						Driver Age
	BTW Total Points Off	Rapid Pace Walk Time	Trails A Time	Trails B Time	Maze 1 Time	Maze 2 Time	Mazes 1 + 2 Time	
Percent Trips 8 p.m.-6 a.m.	(188) -0.17	(145) -0.06	(209) 0.11	(208) 0.01	(187) -0.09	(190) -0.06	(187) -0.10	(231) 0.05
Percent Trips Under 1 mile	(187) -0.01	(144) 0.08	(208) 0.02	(207) 0.03	(186) 0.04	(189) 0.08	(186) 0.07	(231) ***0.20***
Percent Trips 1-2.5 miles	(187) 0.15	(144) 0.15	(208) -0.00	(207) 0.07	(186) -0.11	(189) 0.08	(186) -0.05	(231) -0.12
Percent Trips 2.5-5 miles	(187) 0.14	(144) 0.02	(208) 0.09	(207) 0.04	(186) -0.05	(189) 0.06	(186) -0.01	(231) -0.02
Percent Trips 5-10 miles	(187) -0.10	(144) -0.09	(208) -0.13	(207) -0.12	(186) 0.02	(189) -0.09	(186) -0.03	(231) -0.10
Percent Trips 10-20 miles	(187) ***-0.20***	(144) -0.07	(208) 0.06	(207) -0.01	(186) 0.08	(189) -0.09	(186) 0.02	(231) 0.04
Percent Trips Over 20 miles	(187) -0.02	(144) -0.18	(208) -0.04	(207) -0.04	(186) 0.03	(189) -0.12	(186) -0.03	(231) -0.07
Percent Trips Night	(188) 0.00	(145) 0.02	(209) -0.06	(208) 0.03	(187) -0.20	(190) -0.06	(187) -0.18	(232) -0.24
Percent Trips Rain/Fog/Wet	(188) -0.03	(145) -0.05	(209) -0.00	(208) -0.05	(187) 0.08	(190) 0.07	(187) 0.09	(232) -0.06
Percent Trips Passengers	(128) ***-0.22***	(67) ---	(130) -0.14	(129) -0.07	(128) -0.03	(130) -0.11	(128) -0.07	(153) 0.04
Percent Trips Rush Hour	(188) -0.10	(145) 0.01	(209) -0.18	(208) -0.15	(187) -0.05	(190) -0.15	(187) -0.11	(231) -0.04
Trip Speed Average (Minimum)	(187) 0.10	(126) -0.02	(189) -0.16	(188) -0.02	(186) -0.08	(189) -0.03	(186) -0.08	(212) -0.04
Trip Speed Average (Maximum)	(187) -0.03	(126) ***-0.20***	(189) ***-0.26***	(188) ***-0.21***	(186) -0.13	(189) -0.15	(186) -0.17	(212) ***-0.20***
Trip Speed Average (Mean)	(187) -0.03	(126) -0.14	(189) -0.17	(188) -0.12	(186) -0.06	(189) -0.16	(186) -0.12	(212) -0.15

Note: Bold text indicates weak (as opposed to very weak) correlations.

SHRP2 Sample

Planned Analyses

Under a data use license with the SHRP2 data custodian, researchers obtained a sample of 1,096 drivers age 60 and older. Data for three functional measures (TMT A, TMT B, and Rapid Pace walk) were missing for 102 drivers; these were excluded from the analysis sample. An additional 12 older SHRP2 participants were excluded because their cumulative days of vehicle instrumentation or trip counts as individuals were 0.01% or lower in the distribution of cumulative total vehicle instrumentation days or total trip counts for all older participants—i.e., those older SHRP2 drivers who very rarely drove. Following are summary statistics for the remaining sample pertaining to driver performance and exposure variables, driver functional status, and the correlations between them. The final SHRP2 analysis data set was 982 participants, including 455 females and 527 males ranging from 60 to 98 years old, who made a total of 1,586,210 trips, driving for a total of 357,856 hours and 9,749,341 miles. Table 4 presents summary statistics for this sample.

Table 4: Summary Statistics for Variables Analyzed in the SHRP2 Data

Variable	N	Min.	Max.	Mean	SD	Median
Driver Age	982	60.0	98.0	74.3	7.3	75.0
Functional Ability Measures						
Rapid Pace Walk Time (seconds)	962	2.7	15.7	6.1	1.8	5.8
Trails A Time (seconds)	977	13.7	326.6	40.5	18.3	37.0
Trails B Time (seconds)	969	36.9	359.7	102.8	43.7	94.3
Driving Exposure						
Trip Count	982	42	9,946	1,615.3	1,448.2	1,222.5
Vehicle Equipment Days	982	19	1,129	504.9	223.6	420.0
Shortest Trip Time (minutes)	982	1.0	5.7	1.1	0.2	1.0
Longest Trip Time (minutes)	982	11.3	355.4	137.9	68.8	125.8
Mean Trip Time (minutes)	982	2.2	47.7	13.6	4.5	12.8
Total Driving Time (minutes)	982	500.4	222,227.9	21,864.9	23,406.5	15,618.7
Shortest Trip Distance (miles)	982	0.1	0.4	0.1	0.0	0.1
Longest Trip Distance (miles)	982	5.2	374.3	114.0	78.7	94.4
Mean Trip Distance (miles)	982	0.8	26.9	6.3	3.3	5.6
Total Driving Distance (miles)	982	130.2	76,321.3	9,928.0	10,672.6	6,982.1
Trip Count per Day	982	0.1	13.0	3.1	1.9	2.8
Minutes Driven per Day	982	0.0	293.0	41.5	31.2	35.0
Miles Driven per Day	982	0.0	84.5	17.8	14.3	14.3
Maximum Speed (mph)	982	48.5	93.2	80.5	7.6	81.1
Percent Trips 6 a.m. – 10 a.m.	982	0.2	59.2	15.5	9.9	13.3
Percent Trips 10 a.m. – 3 p.m.	982	11.7	90.8	49.2	12.9	48.9
Percent Trips 3 p.m. – 8 p.m.	982	1.7	80.1	29.9	11.1	30.0
Percent Trips 8 p.m. – 6 a.m.	982	0.0	55.9	5.4	5.3	4.1

Variable	N	Min.	Max.	Mean	SD	Median
Percent Trips Under 1 mile	982	1.1	90.4	18.7	11.4	16.0
Percent Trips 1 – 2.5 miles	982	3.6	63.0	24.9	12.0	22.9
Percent Trips 2.5 – 5 miles	982	0.7	69.8	23.6	10.9	22.7
Percent Trips 5 – 10 miles	982	0.4	67.7	18.2	10.9	16.2
Percent Trips 10 – 20 miles	982	0.0	58.1	9.5	8.1	7.1
Percent Trips Over 20 miles	982	0.0	38.7	5.2	6.0	3.2
Percent Trip Night	982	0.0	60.5	6.4	6.1	4.8
Percent Trips Rush Hour	982	6.1	77.4	36.8	11.3	36.1
Trip Speed Average (Minimum)	982	3.2	13.9	6.2	1.4	6.0
Trip Speed Average (Maximum)	982	25.6	80.4	60.7	10.0	61.7
Trip Speed Average (Mean)	982	13.9	41.4	27.3	4.3	27.2

Within the analysis sample, isolated measures of functional ability were missing for 38 of the 982 participants—Rapid Pace Walk (missing for 20 participants), TMT A (missing for 5), Trails B (missing for 12), and TMT A and TMT B (missing for 1)—but these participants were retained. The analyst performed correlations for each of the exposure measures with each functional ability measure. Using the same classification system as cited above, most of the correlations between functional ability and exposure were very weak with two exceptions that were "weak;" these are highlighted in **bold** in Table 5.

Table 5: Correlations Between Age, Functional Ability, and Driving Exposure

Driver Age and Driving Exposure Measures	**Functional Ability Measures**			**Driver Age** (n=982)
	Rapid Pace Walk Time (n=962)	Trails A Time (n=977)	Trails B Time (n=969)	
Driver Age	***0.26***	***0.28***	***0.29***	----
Shortest Trip Time	0.08	0.04	0.04	0.08
Longest Trip Time	-0.10	-0.02	-0.12	-0.18
Average Trip Time	-0.05	-0.05	-0.08	***-0.22***
Total Driving Time	0.07	0.13	0.11	-0.10
Shortest Trip Distance	-0.01	-0.03	-0.04	-0.07
Longest Trip Distance	-0.18	-0.06	-0.19	***-0.22***
Mean Trip Distance	***-0.21***	-0.10	-0.17	***-0.29***
Total Driving Distance	-0.04	0.09	0.01	-0.15
Trip Count per Day	0.06	0.10	0.10	-0.04
Minutes Driven per Day	0.05	0.07	0.07	-0.15
Miles Driven per Day	-0.10	0.01	-0.05	***-0.22***
Maximum Speed	-0.10	0.04	-0.05	-0.09
Percent Trips 6 a.m. – 10 a.m.	-0.12	-0.05	-0.06	-0.08
Percent Trips 10 a.m. – 3 p.m.	0.13	0.07	0.12	***0.26***
Percent Trips 3 p.m. – 8 p.m.	-0.01	-0.01	-0.06	-0.10
Percent Trips 8 p.m. – 6 a.m.	-0.08	-0.05	-0.05	***-0.26***
Percent Trips Under 1 mile	0.16	0.08	0.08	0.18
Percent Trips 1-2.5 miles	0.04	0.05	0.10	0.14
Percent Trips 2.5-5 miles	0.03	0.04	0.00	0.08
Percent Trips 5-10 miles	-0.07	-0.05	-0.06	-0.14
Percent Trips 10-20 miles	-0.12	-0.11	-0.08	***-0.20***
Percent Trips Over 20 miles	-0.16	-0.07	-0.11	***-0.24***
Percent Trips Night	-0.12	-0.05	-0.06	***-0.28***

Driver Age and Driving Exposure Measures	Functional Ability Measures			Driver Age (n=982)
	Rapid Pace Walk Time (n=962)	Trails A Time (n=977)	Trails B Time (n=969)	
Percent Trips Rush Hour	-0.13	-0.09	-0.12	***-0.22***
Trip Speed Average (Minimum)	-0.13	-0.16	-0.15	***-0.26***
Trip Speed Average (Maximum)	-0.19	-0.07	-0.19	***-0.26***
Trip Speed Average (Mean)	***-0.27***	-0.15	-0.17	***-0.39***

Exploratory Analyses

Exploratory analyses focused on event-based response latencies, hypothesizing significant effects related to driver cognitive status. Initially, researchers hypothesized that older SHRP2 participants with a serious cognitive impairment on at least one cognitive test would have longer latencies in crash and/or near-crash events than those without serious cognitive impairment. Response latency was broken into two parts. "Latency 1" began the instant a driver *could* perceive a "precipitating event," based on video review by the data custodian's analyst, and ended when the driver began to react, i.e., when it was apparent that the driver *had* perceived the threat and understood the need for a control movement. "Latency 2" began with the driver's control movement and ended at the time of impact/closest proximity. Scores on each of the three cognitive tests (TMT B, VMI, and UFOV) were categorized as either *no serious impairment* (combining no impairment and mild impairment) or *serious impairment* as defined in terms of these cut points: TMT B: ≥ 180 seconds; VMI: ≥ 5 errors; and UFOV: ≥ 300 ms. Cut point values reflect test scores associated with the maximum valid odds ratios obtained for each of these significant predictors of older driver crash risk (Staplin et al., 2003). Researchers selected event types and relied on fault status (as determined by the data custodian's video coders) to define eight candidate analysis data sets, "A" through "H," as shown in Table 6.

Table 6: Candidate Data Sets for NDS Event-Based Latency Analyses

Analysis Set	Number of Drivers	Event Type	Fault Status
A	487	Crash or Near-Crash	At Fault or Not at Fault
B	51	Crash Only	At Fault or Not at Fault
C	469	Near-Crash Only	At Fault or Not at Fault
D	338	Crash or Near-Crash	At Fault
E	31	Crash Only	At Fault
F	327	Near-Crash Only	At Fault
G	15	Rear-End, Striking - Crash Only	At Fault or Not at Fault
H	75	Rear-End, Striking, Near-Crash Only	Not at Fault

Perception-Reaction Time Latency 1 Analysis. SHRP2 data discriminate two components of a driver's perception-reaction time (PRT). "Latency 1" corresponds to the time in milliseconds from an event start to the driver's reaction start; "Latency 2" corresponds to the time in milliseconds from the driver's reaction start to time of impact (or time of closest proximity, if a near-crash). Table 7 shows the results of *t*-tests for significant differences between participants with and without serious cognitive impairments. As highlighted, two comparisons were statistically significant at the 0.10 level but not at the 0.05 level of significance, both included at-fault *and* not-at-fault drivers. Note that the sample sizes were smaller than those shown above, due to missing latencies for some participants. There were too few drivers for a valid test for analysis set "G."

Table 7: PRT Latency 1 Analysis Results

Analysis Set	Event Type	Fault	Outcome Measure = Latency 1: Time From Event Start to Driver RT (ms)				Significance Level One-Tailed t-Test
			No Serious Cognitive Impairment		Serious Cognitive Impairment		
			N	Mean	n	Mean	
A	Crash or Near-Crash	Regardless of Fault	308	1.50	163	1.70	$t(469) = -1.55$, p = 0.06
B	Crashes Only	Regardless of Fault	22	2.23	12	2.59	$t(32) = -0.58$, p = 0.28
C	Near-Crashes Only	Regardless of Fault	297	1.49	161	1.88	$t(456) = -1.56$, p = 0.06
D	Crash or Near-Crash	At Fault	207	1.95	115	2.03	$t(320) = -0.49$, p = 0.31
E	Crashes Only	At Fault	15	2.44	9	2.47	$t(22) = -0.04$, p = 0.48
F	Near-Crashes Only	At Fault	202	1.94	115	2.01	$t(315) = -0.41$, p = 0.34
H	Rear-End Striking–Near-Crash	Not at Fault	47	1.34	26	1.28	$t(71) = 0.20$, p = 0.42

To further explore the relationships between functional scores and events, subsequent analysis focused on fault status. Specifically, the t-test outcomes that showed nearly identical results for the "crash or near-crash" and "near-crash only" analysis sets could be explained by the fact that there were only nine participants with serious cognitive impairments in at-fault crashes. In other words, in both cases the findings that approached significance could be attributed to the near-crash data. A subsequent Latency 1 analysis of near-crash events demonstrated significant differences within both cognitive status groups as a function of fault status, as shown in Table 8.

Table 8: Latency 1 Analysis Results for Near-Crash Events Only

Event Type	Cognitive Status	Outcome Measure = Latency 1: Time From Event Start to Driver Response (ms)				Significance Level One-Tailed *t*-Test
		Fault Status: Not at Fault		Fault Status: At Fault		
		n	Mean	n	Mean	
Near-Crash	No Serious Impairment	200	1.12	202	1.94	$t(374) = -6.6$, $p < 0.001$
Near-Crash	Serious Impairment	100	1.17	115	2.00	$t(206) = -4.98$, $p < 0.001$

A two-way analysis of variance (ANOVA) then examined main effects and interactions between fault status and cognitive status in near-crashes. As expected, based on the previous within-group tests, the ANOVA showed a significant main effect of fault status on Latency 1 (F = 67.01; df = 1, 613; $p<.0001$); however, there was no main effect of cognitive status, nor a significant interaction between fault status and cognitive status. Thus, there was no significant relationship between cognitive status and Latency 1, regardless of fault status.

PRT Latency 1 Re-analysis. Given the basis for classifying the cognitive status of the older NDS participants – i.e., scores on tests of visual search, visual information processing, and divided attention capabilities – researchers refined their earlier hypothesis to take event type into account. Specifically, the 471 drivers in the crash and near-crash events captured in analysis set "A," the largest analysis set, were sorted by "precipitating event." The SHRP2 event table dictionary describes this as "the state of environment or action that began the event sequence under analysis. What environmental state or what action by the subject vehicle, another vehicle, person, animal, or non-fixed object was critical to this vehicle becoming involved in the crash or near-crash" (Hankey et al., 2016).

In this reanalysis the planned test of the hypothesis that drivers with a serious cognitive impairment would demonstrate a significantly longer Latency 1 was limited to only those crashes and near-crashes in which the precipitating event involved a peripheral, as opposed to a focal, threat. This follows from research showing that compared to younger drivers, older drivers scan left and right less frequently and are more likely to focus straight ahead or in the intended direction of travel (Romoser et al., 2013; Dukic & Broberg, 2012; Bao & Boyle, 2009; Scott et al., 2013). This behavior may explain older drivers' over-involvement in angle crashes at intersections as the struck vehicle, as well as their overrepresentation in "looked but did not see" crashes and being charged with "failure to yield" when involved in crashes with other vehicles (Stutts et al., 2008). Logically, drivers who experience relatively greater age-related declines in visual information processing and visual search ability would take longer to detect peripheral (but not focal) threats. The list below presents threats identified by the data custodian's video coders, sorted into those the researchers characterized as "focal" versus "peripheral."

Focal threat-related events

- Animal in roadway
- Object in roadway
- Other type of event not attributed to subject vehicle
- Other vehicle ahead - at a slower constant speed
- Other vehicle ahead - decelerating
- Other vehicle ahead - slowed and stopped 2 seconds or less
- Other vehicle ahead - stopped on roadway more than 2 seconds
- Pedalcyclist or other type of non-motorist in roadway
- Pedestrian in roadway
- Subject over left edge of road
- Subject over left lane line
- Subject over right edge of road
- Subject vehicle - end departure
- Subject vehicle ahead - decelerating
- Subject vehicle backing
- Subject vehicle, other
- This vehicle lost control - excessive speed
- Unknown

Peripheral threat-*related events*

- Animal approaching roadway
- Another vehicle backing

- Another vehicle making U-turn
- Other vehicle traveling in opposite direction
- Other vehicle entering intersection – left turn across path
- Other vehicle entering intersection - right turn across path
- Other vehicle entering intersection - straight across path
- Other vehicle entering intersection - turning onto opposite direction
- Other vehicle entering intersection - turning same direction
- Other vehicle from driveway - turning into opposite direction
- Other vehicle from driveway - turning into same direction
- Other vehicle from parallel or diagonal parking lane
- Other vehicle lane change - left in front of subject
- Other vehicle lane change - left, sideswipe threat
- Other vehicle lane change - right in front of subject
- Other vehicle lane change - right, sideswipe threat
- Other vehicle oncoming - over left line
- Pedalcyclist or other type of non-motorist approaching roadway
- Pedestrian approaching roadway
- Subject in intersection - passing through
- Subject in intersection - turning left
- Subject in intersection - turning right
- Subject lane change - left behind vehicle
- Subject lane change - left in front of vehicle
- Subject lane change - left, sideswipe threat
- Subject lane change - right in front of vehicle
- Subject lane change - right, sideswipe threat
- Subject vehicle making a U-turn

Of the 471 drivers who experienced crashes or near-crashes, 280 each faced a peripheral precipitating event and 191 a focal precipitating event. The researchers performed a *t*-test on the 280 peripheral-threat-related crashes and near-crashes. Participants were grouped based on cognitive status. The cognitive impairment group included those with serious impairment in one of the three tests (UFOV, TMT B, VMI); the mild/no impairment group members had no serious impairments in the functional abilities measured by these three tests. An F-Test indicated unequal variances between groups, therefore a two-sample *t*-test assuming unequal variances was performed. This test found that, as predicted, Latency 1 was significantly longer [$t(165) = -2.10$, p=0.019] for the cognitive impairment group (n=98, mean latency=1.47 seconds) than for the mild/no impairment group (n=182, mean latency =1.16 seconds). Finally, a two-way ANOVA applied to these data showed main effects of fault status (F=15.25, p=0.0001) and cognitive status (F=3.87, p=0.05) with no significant interaction between these variables.

RT Latency 2 Analysis. As noted earlier, the second component of PRT coded by the data custodian – Latency 2 – corresponds to the time in milliseconds from the beginning of the driver's reaction to the time of impact in a crash, or time of closest proximity in a near-crash. Researchers had no basis upon which to predict any relationship between drivers' cognitive status and this measure but, given this measure's focus on drivers' movement-to-control, hypothesized a significant difference in Latency 2 between drivers with and without serious impairment in lower limb strength/mobility. A metric for this functional ability was obtained at the time of SHRP2 enrollment using the Rapid Pace Walk. Drivers with walk times of 9 seconds or longer were categorized as having serious impairment in this functional ability.

These analyses included all events captured in analysis set "A": crashes and near-crashes, without regard to fault, and including both peripheral and focal precipitating events (see Table 6). The sample size was 463 drivers, as walk time data were missing for 8 participants. A two-tailed *t*-test, assuming unequal variances, showed that the latency between driver reaction and impact/closest proximity was significantly shorter for drivers with serious impairment in lower limb mobility (mean = 1.01 seconds) than for those without serious impairment (mean = 1.49 seconds), $t\ (40) = 3.76$, $p<0.0005$. There was no significant difference in Latency 2 as a function of cognitive status.

Discussion and Conclusions

This research sought to examine the relationships between a common set of predictors and measures of driver performance and exposure through aggregation of six data sets across prior NHTSA investigations that used common data collection methodologies. Researchers applied criteria to determine which variables would be included in the synthesis, beginning with the requirement that only variables for which data were collected for *at least half of all participants* in the pooled sample, i.e., for at least 116 of the total of 232 drivers who participated in the prior studies, would be included. This resulted in a single variable for on-road driver performance – the *total error score* on a standardized road test administered by a CDRS. A much broader set of exposure measures obtained through naturalistic observation, as well as measures of functional ability, met the inclusion criterion. Thus, the NHTSA combined data set analyses calculated the correlations between driver age, functional ability, road test score, and an array of trip duration (time), speed, and distance-traveled measures.

The prior NHTSA studies, while sharing common methodologies, also shared certain limitations: Data were collected only in Virginia and the Carolinas, sample sizes were typically small (20 to 60 older drivers), and naturalistic observations of driving behavior were obtained for relatively short intervals (typically 2 to 4 weeks). Recognizing these shortcomings, researchers obtained data from a supplementary source, the SHRP2 data set, that offered sufficient overlap in data collection methods to support parallel analyses and comparison of results with those of the NHTSA data set. The SHRP2 sample included 1,045 drivers 60 and older; observation periods were 1 to 2 years. The computer-based functional assessments performed on driver enrollment in SHRP2 used the identical software platform and many of the same measures as applied in the NHTSA studies.

One additional attribute of the SHRP2 data set was the availability of direct safety measures (crashes and near-crashes) as outcome/criterion variables. None of the SHRP2 kinematic performance data could be unambiguously related to a CDRS driver error score without viewing the video of each event and coding the instantaneous demands of the driving context. Such an effort was beyond the scope of this project. Thus, the SHRP2 data analyses calculated the correlations between age, functional ability, and driving exposure. The research team also took advantage of the crash and near-crash information to test hypotheses that drivers' responses during these events would differ significantly in relation to their functional status.

The combined data set yielded correlations that were overwhelmingly very weak (.00 - .19), with a few exceptions which were weak (.20 - .39). At best, researchers could account for only 10% of the variance in these analyses, and that was for the association between scores on one of the functional (cognitive) status measures and driver age. However, among the isolated weak correlations involving functional status measures, performance, and exposure, the nature of the correlation was consistent with expectations: decreasing function (poorer TMT and Rapid Pace Walk times) was associated with lower maximum average trip speed, and poorer performance on both the TMT and the Maze Test was associated with worse behind-the-wheel performance (higher error scores).

Notwithstanding the findings noted above, the research team was challenged to interpret the NHTSA data set analysis outcomes. Conceivably, a longer period of naturalistic data collection, as provided by the SHRP2 data, might yield stronger and more reliable associations between functional status and measures of exposure if such associations exist. While the lack of stronger relationships between functional status and driving exposure might be attributed to small samples or inadequate data collection intervals in the combined data, similar findings from the SHRP2 data suggest that these results simply indicate that functional status has limited influence on older adults' driving habits.

The weak association between functional status and road test scores was more problematic, and the difficulties may rest with the measurement protocol itself. Although widely regarded as the gold standard for assessing driver performance, a CDRS evaluation is by its nature subjective, making between-evaluator differences likely. Though the same (modified Miller) instrument for scoring driver errors was used throughout, it was applied by three different CDRSs across these NHTSA studies—and this potential confound increases the inevitable variability in driving situational demands from one evaluation to another, even with the same evaluator and test route.

Results of the SHRP2 analysis were informative in several respects. First, the persistent weak relationships between functional status and exposure reinforced the results of the combined data set analyses, which were based on a smaller data set. The interpretation of this finding is consistent with the conclusion expressed in previous NHTSA research that addressed the extent to which generally healthy older adults' choices about how much and under what conditions to drive depends on their level of function: Not a lot (Staplin et al., 2019b). The findings suggest that older adults' choices about when, where, and how often they drive primarily reflect habits and mobility needs, rather than fitness level.

The SHRP2 data analyses also explored how drivers' vehicle control responses in crash and near-crash situations are related to differences in specific functional abilities. At issue was response latency from the instant a driver *could* perceive a "precipitating event," based on video review by the data custodian's analyst, to the driver's reaction start, i.e., when it was apparent that the driver *had* perceived the threat and understood the need for a control movement—termed "Latency 1." Also of interest was the interval beginning with the driver's control movement and ending at the time of impact/closest proximity—termed "Latency 2." An initial test of the hypothesis that drivers with no serious cognitive impairment would demonstrate briefer Latency 1 measures, on average, than those with serious cognitive impairment achieved statistical significance at the 0.10 level but not at the 0.05 level of significance. Researchers then sorted precipitating events into those involving peripheral, as opposed to focal threats, given the specific measurement constructs targeted by the Useful Field of View test (visual information processing speed with divided attention) and the TMT Part B (visual search with divided attention). A re-analysis testing the same hypothesis but including only peripheral-threat-related crashes and near-crashes, based on previous research showing older drivers tend to scan left and right less often than other drivers, demonstrated an average Latency 1 interval for the seriously cognitively impaired SHRP2 drivers that was statistically (p=0.019) longer—by a full three-tenths of a second.

The analysis of the Latency 2 data yielded results that at first seemed counterintuitive: the time between driver reaction and impact was, on average, significantly shorter for drivers *with* a serious impairment in lower limb mobility than for those *without* a serious impairment. To the extent such an impairment resulted in the vehicle travelling farther toward the point of impact before the driver reacted with a control movement; however, it makes sense that it would decrease the distance—and time—associated with that interval.

The SHRP2 analysis results confirm observations of older driver behavior/habits demonstrated in smaller, cross-sectional studies sponsored by NHTSA and others. The analysis results extend beyond those findings to affirm researchers' models of how age-related functional changes (particularly cognition) may affect driving performance. The re-analysis of SHRP2 response latency data specific to precipitating events in the periphery highlights the need to frame research questions precisely in future studies seeking to understand how the capabilities and limitations of normally aging adults may influence the crash avoidance behaviors of these drivers.

References

Bao, S., & Boyle, L. N. (2009). Age-related differences in visual scanning at median-divided highway intersections in rural areas. *Accident Analysis and Prevention*, *41,* 146-152. https://doi.org/10.1016/j.aap.2008.10.007

Dukic, D., & Broberg, T. (2012). Older drivers' visual search behavior at intersections. *Transportation Research Part F, 15,* 462-470. https://doi.org/10.1016/j.trf.2011.10.001

Deng, H., & Wickham, H. (2011, September). Density estimation in R [Electronic publication]. https://vita.had.co.nz/papers/density-estimation.pdf

Evans, J. D. (1996). Straightforward statistics for the behavioral sciences. Brooks/Cole Publishing.

Hankey, J. M., Perez, M. A., & McClafferty, J. A. (2016). Description of the SHRP 2 naturalistic database and the crash, near-crash, and baseline data sets. Virginia Tech Transportation Institute. https://vtechworks.lib.vt.edu/bitstream/handle/10919/70850/SHRP_2_CrashNearCrashBaselineReport_4-25-16.pdf?sequence=1

Joyce, J., Lococo, K. H., Gish, K. W., Mastromatto, T., Stutts, J., Thomas, D., & Blomberg, R. (2018, April). *Older driver compliance with license restrictions* (Report No. DOT HS 812 486). National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/49163

National Center for Statistics and Analysis. (2020, April). *Older population: 2018 data* (Traffic Safety Facts. Report No. DOT HS 812 928). National Highway Traffic Safety Administration. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812928.

Romoser, M., Pollatsek, A., Fisher, D. L., & Williams, C. C. (2013). Comparing glance patterns of older versus younger experienced drivers: scanning for hazards while approaching and entering the intersection. *Transportation Research Part F, 16,* 104-116. https://doi.org/10.1016/j.trf.2012.08.004

Scott, H., Hall, L., Litchfield, D., & Westwood, D. (2013). Visual information search in simulated junction negotiation: gaze transitions of young novice, young experienced, and older experienced drivers. *Journal of Safety Research*, *45,* 111-116. https://doi.org/10.1016/j.jsr.2013.01.004

Staplin, L., Gish, K., & Wagner, E. (2003). MaryPODS revisited: Updated crash analysis and implications for screening program implementation. *Journal of Safety Research, 34*(4)*,* 389-397. https://doi.org/10.1016/j.jsr.2003.09.002

Staplin, L., Lococo, K. H., Gish, K. W., Stutts, J., & Srinivasan, R. (2019a, September). *Activity level, performance and exposure among older drivers* (Report No. DOT HS 812 734). National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/42285

Staplin, L., Lococo, K. H., Mastromatto, T., Gish, K. W., Golembiewski, G., & Sifrit, K. J. (2019b, January). *Mild cognitive impairment and driving performance* (Report No. DOT HS 812 577). National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/38688

Staplin, L., Lococo, K. H., Stutts, J., Robison, K., Levitt, D., Srinivasan, R., & Sifrit, K. J. (in press). *Physical fitness training and older driver performance and exposure*. National Highway Traffic Safety Administration.

Staplin, L., Mastromatto, T., Lococo, K. H., Kenneth W. Gish, K. W., & Brooks, J. O. (2017, August). *The effects of medical conditions on driving performance* (Report No. DOT HS 812 439). National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/34990

Stutts, J., Martell, C., & Staplin, L. (2009, June). *Identifying behaviors and situations associated with increased crash risk for older drivers* (Report No. DOT HS 811 093). National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/1880

Thomas, F. D., Graham, L. A., Finstad, K. A., Wright, T. J., Blomberg, R. D., Lococo, K., Gish, K., Staplin, L., Stutts, J., Wilkins, J., Crompton, C., & Sifrit, K. J. (2020, April). *Older drivers' self-regulation and exposure* (Report No. DOT HS 812 930). National Highway Traffic Safety Administration. https://rosap.ntl.bts.gov/view/dot/49163

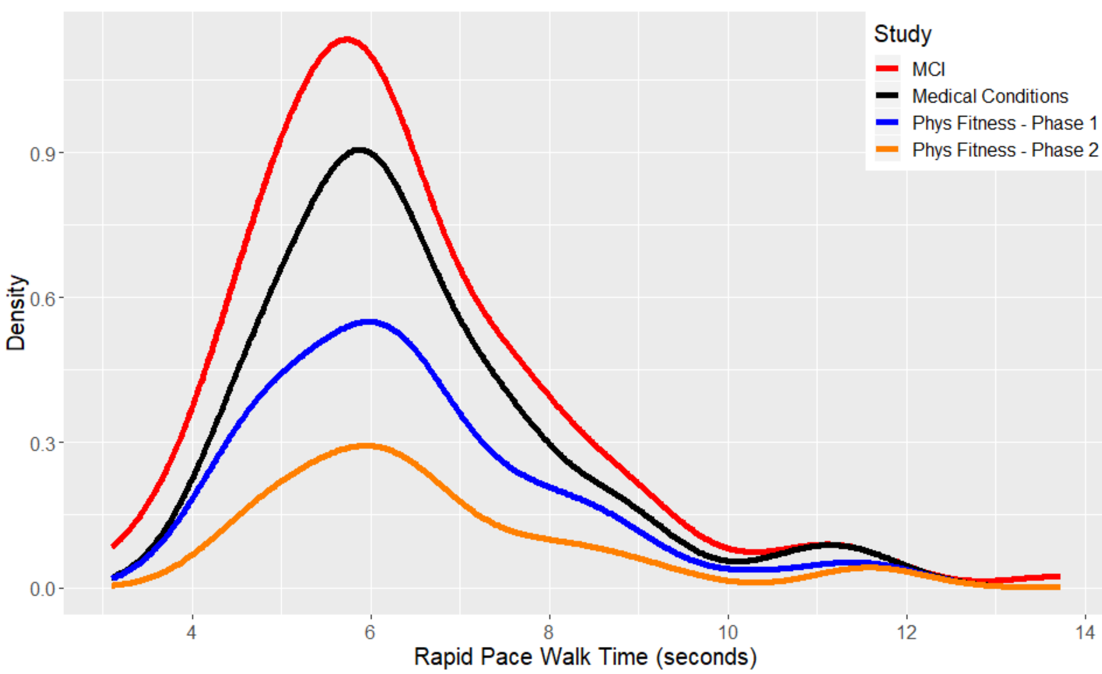
Appendix A: Data Distribution for Clinical and On-Road Performance Measures in NHTSA Studies

The plots in this appendix were generated using the non-parametric Kernel Density Estimation (Deng & Wickham, 2011) function in the R statistical package; each of these may be viewed as a "smoothed histogram" that displays the distribution of scores on each measure as probabilities that a given a score for the variable falls at each of a specified number of equally spaced values under the curve. This technique is useful for comparisons across studies where sample sizes and frequencies of responses differ from study to study. It is more revealing than frequency histograms, which present a dramatically different picture of the data distributions depending upon the selected bin size and are impractical when many values have a zero incidence as is the case with the timed measures.

Rapid Pace Walk

Table A1: Rapid Pace Walk Time (seconds)

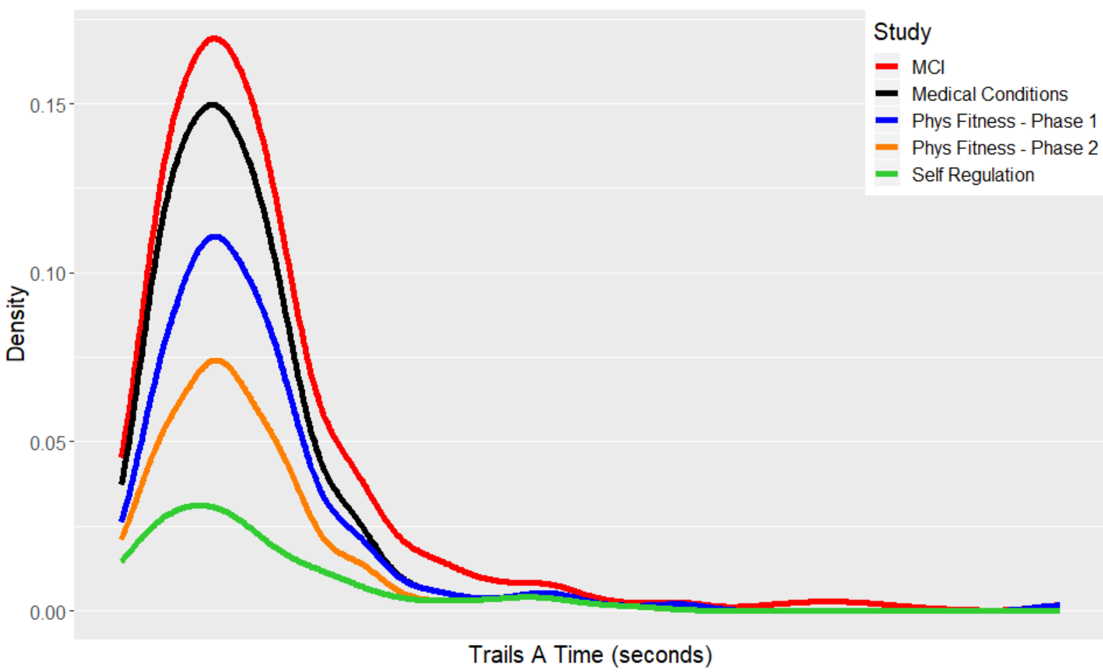
Study	n	mean	median	mode	std	min	max
MCI (Staplin et al., 2019b)	29	6.1	5.5	8.0	2.2	3.1	13.7
Medical Conditions (Staplin et al., 2017)	18	6.5	5.9	4.6	1.6	4.6	11.0
Phys Fitness - Phase 1 (Staplin et al., 2019a)	67	6.3	6.1	5.6	1.6	3.9	11.2
Phys Fitness - Phase 2 (Staplin et al., in press) pre-treatment	31	6.6	6.3	5.6	1.8	4.6	11.7



TMT A

Table A2: TMT A Time (seconds)

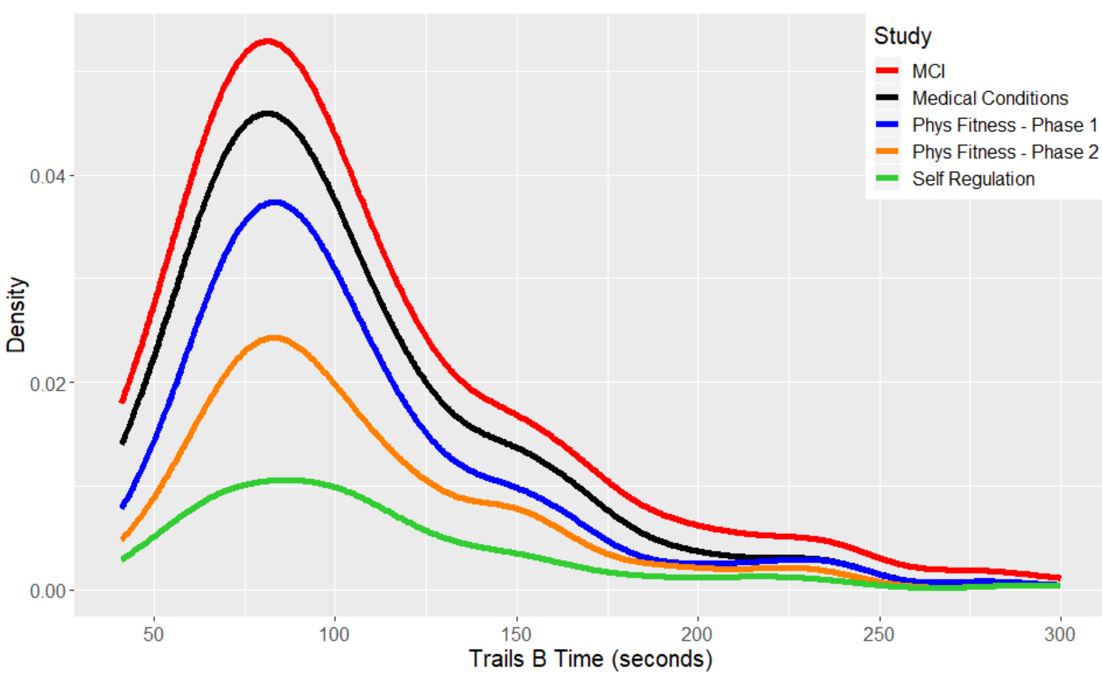
Study	n	mean	median	mode	std	min	max
MCI (Staplin et al., 2019b)	29	50.4	44.2	19.0	27.0	19.0	130.0
Medical Conditions (Staplin et al., 2017)	19	32.7	33.0	27.0	8.2	21.0	52.0
Phys Fitness - Phase 1 (Staplin et al., 2019a)	66	40.5	37.1	25.1	21.0	21.7	164.0
Phys Fitness - Phase 2 (Staplin et al., in press) pre-treatment	31	35.3	33.5	21.5	8.5	21.5	55.8
Self-Regulation (Thomas et al., 2020)	64	38.8	33.5	31.8	18.0	18.0	96.8



TMT B

Table A3: TMT B Time (seconds)

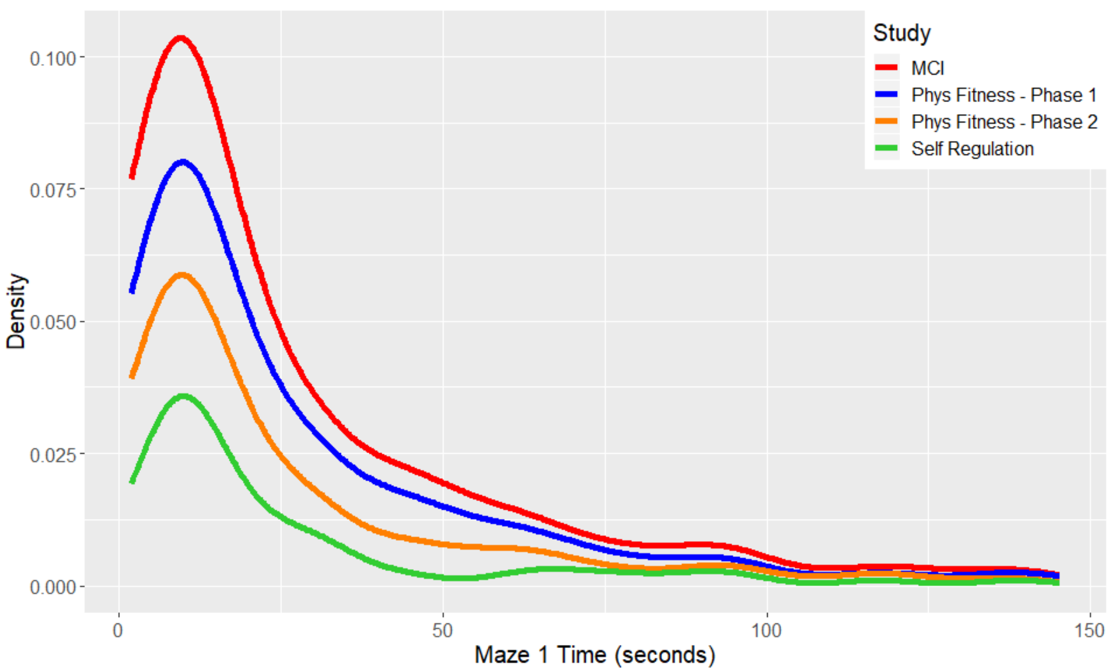
Study	n	mean	median	mode	std	min	max
MCI (Staplin et al., 2019b)	29	125.0	92.1	52.5	68.6	52.5	300.0
Medical Conditions (Staplin et al., 2017)	19	95.4	77.0	44.0	42.5	44.0	184.0
Phys Fitness - Phase 1 (Staplin et al., 2019a)	65	101.0	90.2	41.1	46.5	41.1	275.1
Phys Fitness - Phase 2 (Staplin et al., in press) pre-treatment	31	105.0	91.9	44.6	41.3	44.6	230.0
Self-Regulation (Thomas et al., 2020)	64	109.0	97.5	43.1	49.4	43.1	291.8



Maze 1

Table A4: Maze 1 Completion Time (seconds)

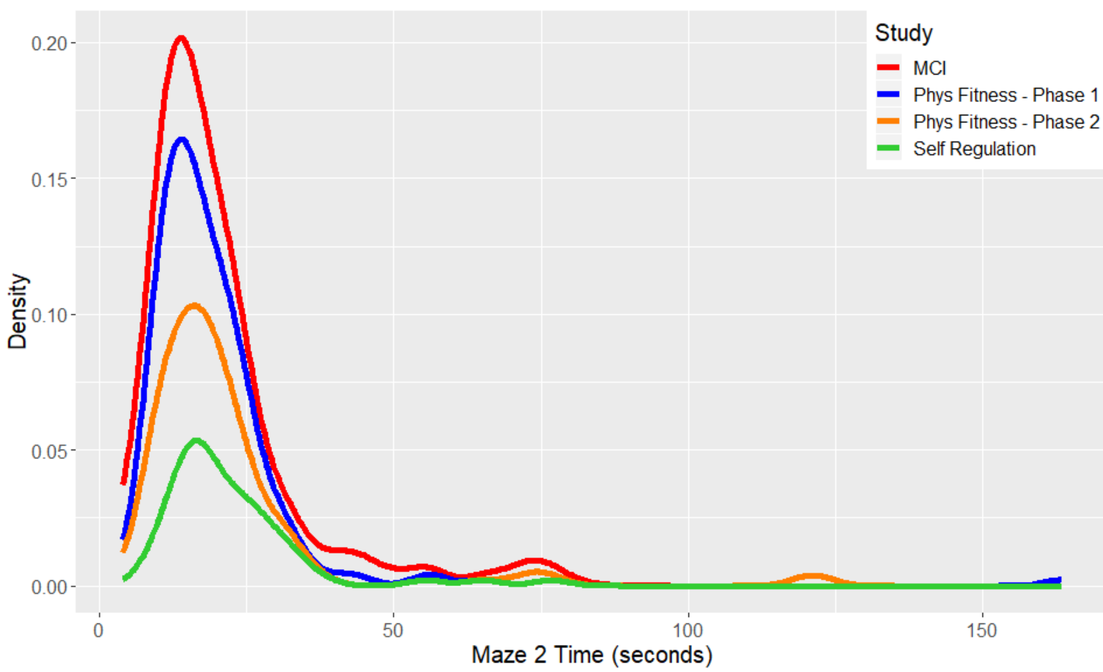
Study	n	mean	median	mode	std	min	max
MCI (Staplin et al., 2019b)	29	24.5	9.5	3.0	31.8	2.0	124.5
Phys Fitness - Phase 1 (Staplin et al., 2019a)	64	29.5	17.0	6.4	30.4	4.1	145.3
Phys Fitness - Phase 2 (Staplin et al., in press) pre-treatment)	30	24.8	9.6	4.3	29.2	4.3	121.9
Self-Regulation (Thomas et al., 2020)	64	26.4	12.0	3.7	29.6	3.7	138.6



Maze 2

Table A5: Maze 2 Completion Time (seconds)

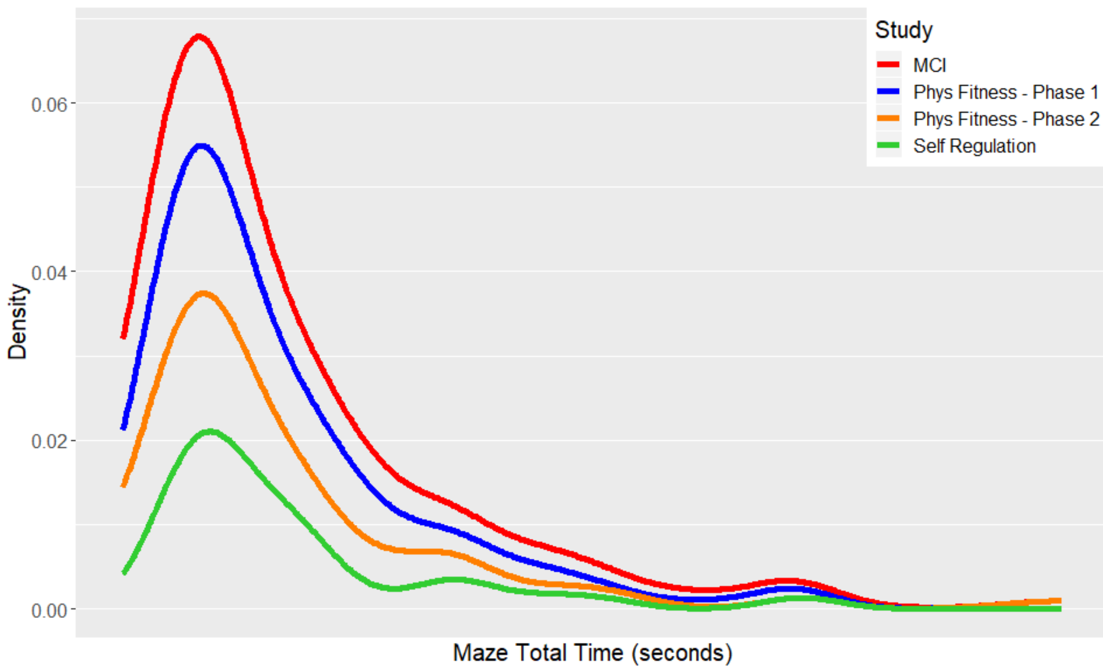
Study	n	mean	median	mode	std	min	max
MCI (Staplin et al., 2019b)	29	22.4	15.4	5.0	18.9	4.0	75.7
Phys Fitness - Phase 1 (Staplin et al., 2019a)	66	19.7	15.2	11.0	20.0	7.3	163.0
Phys Fitness - Phase 2 (Staplin et al., in press) pre-treatment)	30	21.2	16.0	8.2	22.3	8.2	121.0
Self-Regulation (Thomas et al., 2020)	64	22.2	19.4	7.5	12.0	7.5	76.8



Maze Both

Table A6: Maze Total Completion Time (seconds)

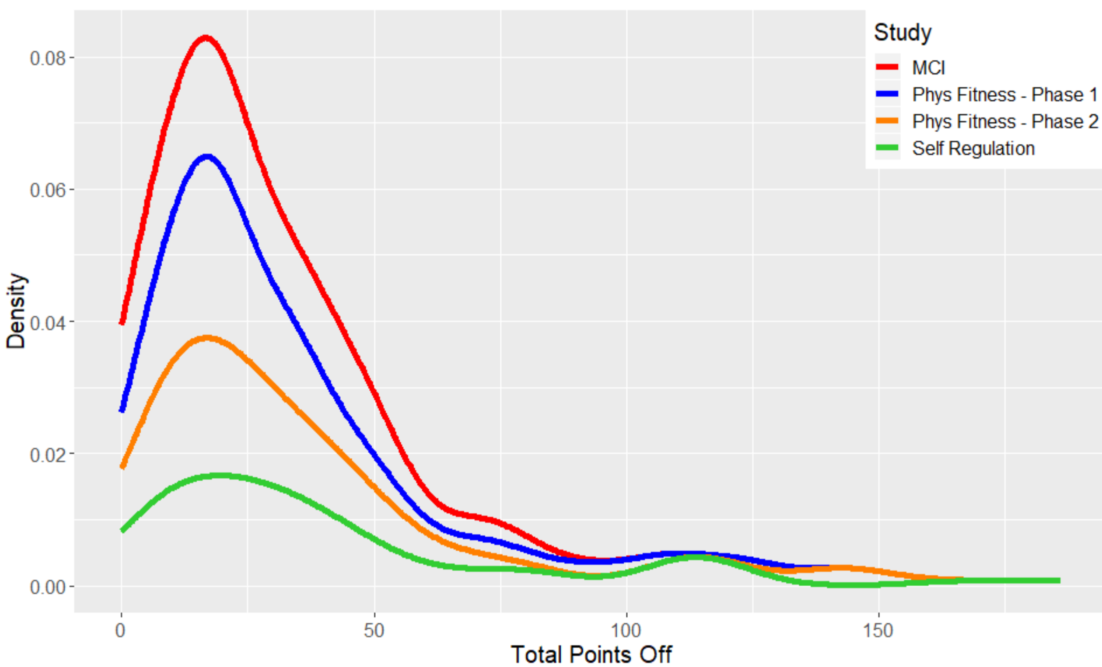
Study	n	mean	median	mode	std	min	max
MCI (Staplin et al., 2019b)	29	47.0	26.2	8.0	44.1	7.0	176.4
Phys Fitness - Phase 1 (Staplin et al., 2019a)	64	49.1	34.3	11.4	37.0	11.4	172.2
Phys Fitness - Phase 2 (Staplin et al., in press) pre-treatment)	30	46.1	29.9	12.5	46.8	12.5	243.1
Self-Regulation (Thomas et al., 2020)	64	48.6	35.9	11.1	35.8	11.1	182.4



Total Points Off Road Test

Table A7: Road Test Score: Total Points Off

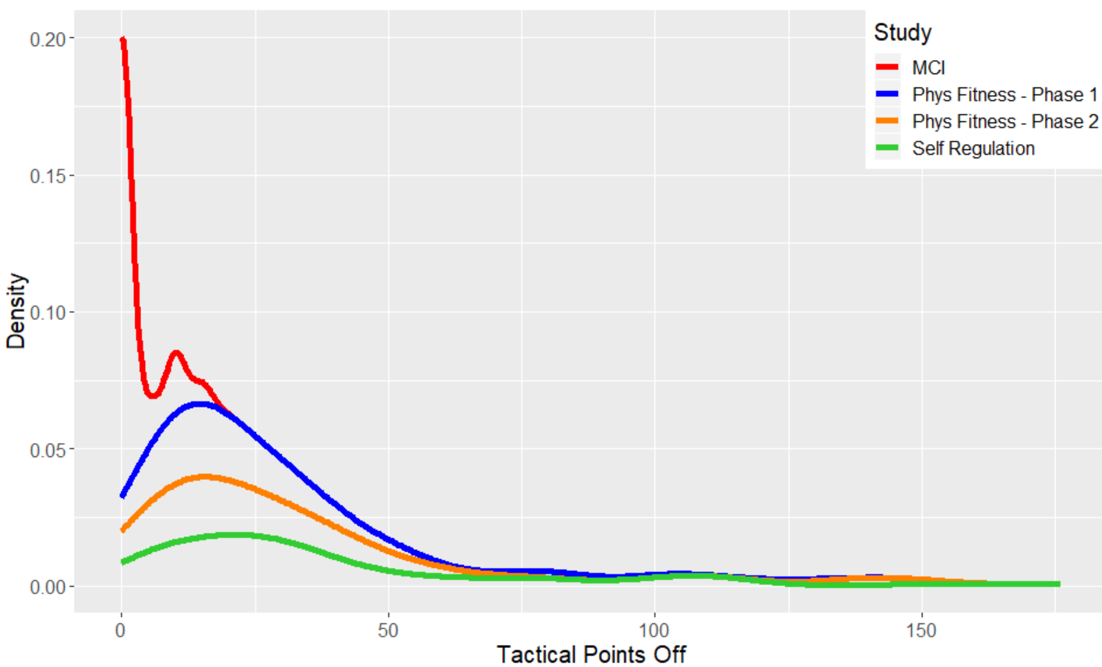
Study	n	mean	median	mode	std	min	max
MCI (Staplin et al., 2019b)	29	25.1	20.0	0	21.0	0	73
Phys Fitness - Phase 1 (Staplin et al., 2019a)	64	29.3	20.0	16	26.1	0	126
Phys Fitness - Phase 2 (Staplin et al., in press) pre-treatment)	31	33.4	23.0	10	33.7	5	143
Self-Regulation (Thomas et al., 2020)	64	42.5	29.5	5	40.6	5	186



Tactical Points Off Road Test

Table A8: Road Test Score: Tactical Points Off

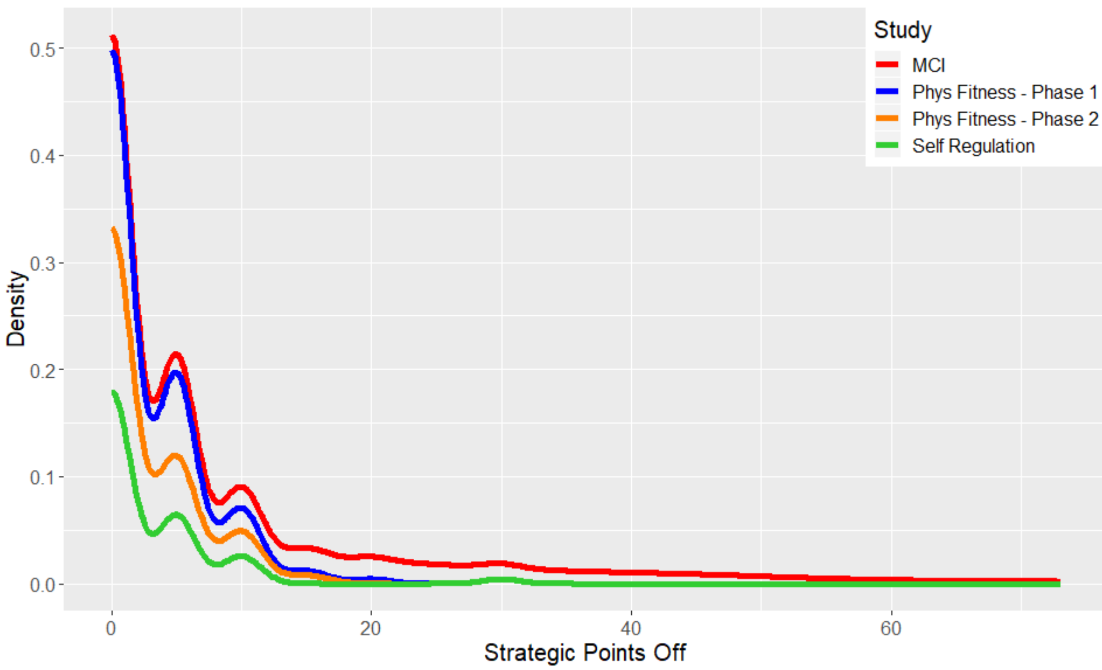
Study	n	mean	median	mode	std	min	max
MCI (Staplin et al., 2019b)	29	1.9	0.0	0	4.1	0	15
Phys Fitness - Phase 1 (Staplin et al., 2019a)	64	25.3	18.0	8	24.1	0	126
Phys Fitness - Phase 2 (Staplin et al., in press) pre-treatment	31	30.4	20.0	10	33.1	0	143
Self-Regulation (Thomas et al., 2020)	64	39.4	25.5	5	37.5	5	176



Strategic Points Off Road Test

Table A9: Road Test Score: Strategic Points Off

Study	n	mean	median	mode	std	min	max
MCI	29	23.0	20	0	19.0	0	73
Phys Fitness - Phase 1	64	2.7	0	0	4.2	0	20
Phys Fitness - Phase 2 (pre-treatment)	31	2.6	0	0	4.1	0	15
Self-Regulation	64	2.6	0	0	4.8	0	30



DOT HS 813 181
November 2021

U.S. Department
of Transportation

National Highway
Traffic Safety
Administration

NHTSA

15343-110221-v3