DEVELOPING A YOUNGER DRIVER ASSESSMENT TOOL TECHNICAL MEMORANDUM: PHASE 1 BETA TEST RESULTS

August 2021

Rebecca Brewster
President and Chief Operating Officer
American Transportation Research Institute
Atlanta, GA

Dan Murray
Senior Vice President
American Transportation Research Institute
Minneapolis, MN

Monica M. Luciana, Ph.D.
Professor
University of Minnesota Department of Psychology
Minneapolis, MN

950 N. Glebe Road, Suite 210
Arlington, Virginia 22203
TruckingResearch.org

ATRI RESEARCH ADVISORY COMMITTEE

ACKNOWLEDGEMENT

ATRI would like to thank Driver iQ for their technical support of the Phase I beta test.

ACRONYMS

ANT-S	Attentional Network Test-Short Form
ATA	American Trucking Associations
ATRI	American Transportation Research Institute
AUDIT	Alcohol Use Disorders Identification Test
BFI-II	Big Five Inventory-II
BMI	Body Mass Index
CDL	Commercial Driver's License
DOT	Department of Transportation
GATS	Great American Trucking Show
GED	General Educational Development
MATS	Mid-America Trucking Show
MSIT	Multi-Source Interference Task
MVR	Motor Vehicle Record
PSP	Pre-Employment Screening Program
PSQI	Pittsburgh Sleep Quality Index
RAC	Research Advisory Committee
SOPT	Self-Ordered Pointing Test
SSS-V	Sensation Seeking Scale-V
UPPS-P	Urgency, Premeditation, Perseverance, Sensation Seeking, and Positive Urgency Impulsivity Behavioral Scale
WASI-II	Wechsler Abbreviated Scale of Intelligence
WHO	World Health Organization

TABLE OF CONTENTS

TABLES

INTRODUCTION

Recruiting and retaining sufficient numbers of qualified truck drivers consistently ranks as a top trucking industry concern.[1] Estimates from the American Trucking Associations (ATA) have the driver shortage topping 100,000 by the year 2023 due to projected freight growth, industry retirements and competition from other industries.[2] According to data from the Bureau of Labor Statistics, nearly a third of the industry's workforce (27.4%) is over the age of 55.[3] The aging demographics of the trucking industry's workforce puts significant pressure on the industry to increase the available pool of qualified truck drivers.

One of the issues motor carriers face in finding and keeping drivers is the challenging lifestyle inherent in a long-haul driving career. The job can be stressful with extended time away from home and family, irregular schedules often exacerbated by traffic congestion and customer detention, and the lack of available, safe places to park a truck for a driver's mandated rest breaks.

Another driver recruitment challenge faced by the industry is the federal requirement to be 21 years of age before obtaining a Commercial Driver's License (CDL) in order to operate across state lines. This creates a three-year gap following high school during which potential new entrant drivers seek employment in other industries. The DRIVE-Safe Act, first introduced in Congress in 2018 and reintroduced in 2021, would provide an avenue for 18-20 year olds to drive in interstate operations. Many in the trucking industry view this legislation as a clear pathway to safely integrate younger drivers into trucking careers.[4]

However, one of the concerns with opening the labor pool to younger drivers is that young people engage in numerous driving-related higher risk behaviors. Moreover, the emerging adult period, from 18-25 years of age, is associated with levels of immature cognitive and socioemotional function and with higher rates of risk-taking behavior relative to other time periods in life. Many forms of adolescent and young adult risk-taking behaviors lead to preventable negative outcomes including motor vehicle crashes. Data from the U.S. Department of Transportation (U.S. DOT) indicate that individuals under 24 years of age represented 18.4 percent of drivers involved in fatal crashes in 2017 while representing just 11.8 percent of the licensed driving population that year.[5]

In 2015, the American Transportation Research Institute (ATRI) Research Advisory Committee (RAC)[6] prioritized a top research proposal, submitted by the Minnesota Trucking Association, to develop an assessment tool to identify the safest drivers among 18-20 year olds. The first research task in the Younger Driver Assessment Tool development was the publication of a report, coauthored by Monica Luciana, a faculty member and consultant from the University of

1 "Critical Issues in the Trucking Industry – 2020." American Transportation Research Institute. Arlington, VA. October 2020.
2 Costello, Bob and Alan Karickhoff. (2019). "Truck Driver Shortage Analysis 2019." American Trucking Associations.
3 Current Population Survey (CPS) 2019. U.S. Census Bureau and Bureau of Labor Statistics.
4 Reymer, Jeremy. "Opinion: How the DRIVE-Safe Act Could End the Driver Shortage." Transport Topics. December 3, 2018. Available online: https://www.ttnews.com/articles/opinion-how-drive-safe-act-could-end-driver-shortage
5 "Traffic Safety Facts 2017 Data – Young Drivers." U.S. Department of Transportation National Highway Traffic Safety Administration. May 2019. Available online: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812753
6 ATRI's RAC is comprised of industry stakeholders representing motor carriers, trucking industry suppliers, labor and driver groups, law enforcement, federal government and academia. The RAC is charged with annually recommending a research agenda for the Institute.

Minnesota and an expert in adolescent and young adult neurodevelopment, describing the scientific literature on safe versus unsafe driving. That report concluded that unsafe driving in non-commercial settings could be reliably predicted by a number of individual difference factors, including personality traits (high levels of impulsivity, aggression/frustration, sensation-seeking), health status (high BMI; poor sleep; fatigue, attention deficit disorder), lifestyle factors (increased substance use), and cognitive ability (low intelligence; low executive function). Inexperience was also identified as a significant and salient predictor of driver risk. Many of these factors have been tested as predictors of safety in non-commercial drivers with limited application to the commercial driving population.[7] Thus, it was unclear which of these risk-related characteristics is incrementally predictive of driver safety, above and beyond age alone, in younger drivers and among drivers of commercial vehicles.

Utilizing the findings of that report, ATRI and Dr. Luciana assembled a comprehensive assessment battery to administer to current commercial truck drivers. Truck drivers who participated in the assessment represented a broad range of ages, driving experience and safety performance. This report describes the findings from this Phase 1 Beta Test of the ATRI Younger Driver Assessment Tool.

[7] Boris, Caroline and Monica Luciana. (2017). "Developing a Younger Driver Assessment Tool Technical Memorandum #1." American Transportation Research Institute.

METHODOLOGY

Study Sample

Commercial truck drivers currently employed in the industry were the participants in the Phase 1 Beta Test. The test population was a convenience sample of self-selected drivers recruited at two truck shows and through trucking fleets in Minnesota, Ohio and Pennsylvania. In advance of the 2018 Mid-America Trucking Show (MATS) held in Louisville, Kentucky, ATRI issued a news alert notifying commercial drivers of the opportunity to take part in the assessment while attending the 2018 MATS. Drivers who attended the event were directly informed about the study via advertisements at the ATRI exhibit booth. A similar recruitment strategy was deployed in advance of the Great American Trucking Show (GATS) held in Dallas, Texas in August 2018.

A total of 75 drivers were assessed across the two truck shows. At both MATS and GATS, a large room was reserved by ATRI and set up to accommodate multiple testing stations. Research staff, enabled by laptop computers and other materials, worked individually with enrolled drivers. Following an initial examination of the data collected from the MATS and GATS samples, a third group of drivers (n=19) was recruited through direct contacts with for-hire motor carriers. The goal of this additional recruitment was to increase the study sample size and to increase inclusion of younger drivers. Additional fleets were identified through personal contact with company presidents and/or safety directors of firms that had indicated an interest in having their drivers participate in ATRI's research. These contacts were generally made directly by ATRI or through the participating fleets' State Trucking Associations.

ATRI research staff provided the fleets with information on the study that drivers could review to determine if they were interested in volunteering for the research. ATRI research staff traveled to the trucking companies and set up testing space in a training or conference room at each facility. Drivers willing to participate scheduled an assessment session with trained ATRI staff. There was no direct monitoring of driver behavior (e.g., use of alcohol; medication use) for the immediate period of time prior to testing.

The total combined study sample included 94 drivers who ranged in age from 20 to 60 years. All participants received a $100 gift card for their participation. At enrollment, each driver signed an informed consent form and gave written permission for his/her Motor Vehicle Record (MVR) and Pre-Employment Screening Program (PSP) report, if accessible, to be released to ATRI. These records were used to quantify indices of driver safety. Following the provision of informed consent, drivers were tested on a battery of measures that took approximately two hours to complete.

Data Collection Procedures

Two academic research assistants were recruited and hired to work alongside ATRI staff in data collection. The staff were trained by Dr. Luciana on all study procedures, and observed prior to data collection, to assure a high level of accuracy and standardization of procedures.

Data collection was then directed and monitored by the ATRI staff lead, who secured relevant site licenses for study measures. ATRI staff tabulated the data for statistical analysis and supervised the collection of MVR data. De-identified data files were shared with Dr. Luciana for analysis.

All but two measures (the Demographics Questionnaire and the Matrix Reasoning Task) were administered in automated form using identical laptop computers. The automated measures were administered through a software application whereby computerized questions and experimental tasks were presented to each participant using standardized instructions. Data were logged and scored automatically using algorithms developed for each measure. Data were then available in spreadsheet form for analysis. Computerized assessments were utilized given that they can be administered efficiently in a highly standardized fashion to large numbers of participants, minimizing researcher burden.

Assessment Measures

Demographics questionnaire. A questionnaire was administered that queried age, gender, race/ethnicity, language proficiency, level of education, household income, household and marital status, and characteristics of the driver's employment environment such as years employed, industry sector, fleet size, annual miles driven, typical haul length, age at which the CDL was obtained, and number of years employed in the trucking industry. This questionnaire was not automated and was completed on paper.

Beck Depression Inventory-II. The Beck Depression Inventory II is a 21-item self-report inventory widely used to assess symptoms of clinical depression such as a depressed mood, suicidal thoughts, feelings of guilt or of being punished, and physical symptoms such as changes in sleep, changes in appetite, fatigue, weight gain or loss, and decreased libido.[8] Each question is rated on a four-point scale where a 0 represents no experience with the symptom over the past two weeks and a 3 represents daily or severe levels of the symptom. Total scores can range from values of 0 to 63. Scores from 14-19 suggest mild depression; scores from 20-28 suggest moderate depression; scores above 28 suggest severe depression.

Alcohol Use Disorders Identification Test. The Alcohol Use Disorders Identification Test (AUDIT) is a 10-item self-report screening tool developed by the World Health Organization (WHO) to assess alcohol consumption, drinking behaviors, and alcohol-related problems.[9] Each question is rated on a five-point scale (ranging from 0 to 4). Total scores range from 0 to 40. Total scores above 8 suggest hazardous or problematic alcohol use.[10]

Body Mass Index (BMI). Height and weight were recorded by the research staff and each driver's body mass index was calculated using an automated system maintained by the National Institutes of Health.[11]

Pittsburgh Sleep Quality Index (PSQI). The PSQI is a self-report measure of sleep quality that differentiates "poor" from "good" sleep by measuring seven domains:

- subjective sleep quality;
- sleep latency;

[8] Beck, A.T., Ward, CH, Mendeson, M, Mock, J. & Ergaugh, J. (1961) "An Inventory for Measuring Depressing." *Archives of General Psychiatry*, 6, 561-571; Beck, A.T., Steer, R.A., Ball, R., Ranieri, W. (1996) "Comparison of Beck depression inventories –IA and –II in psychiatric outpatients." *Journal of Personality Assessment*, 67 (3), 588-597.
[9] Babor, T., & World Health Organization. Department of Mental Health Substance Dependence. (2001). "AUDIT, the alcohol use disorders identification test: Guidelines for use in primary care (2nd ed.)." Geneva, Switzerland: World Health Organization, Dept. of Mental Health and Substance Dependence.
[10] Guidelines about use of the AUDIT have been published by the WHO and are available online (http://whqlibdoc.who.int/hq/2001/who_msd_msb_01.6a.pdf).
[11] The calculator is available online at: https://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmicalc.htm

- sleep duration;
- habitual sleep efficiency;
- sleep disturbances;
- use of sleep medication; and
- daytime dysfunction over the last month.[12]

Scores are based on a 0 to 3 scale for each question, where 3 reflects the negative extreme on a Likert Scale. A global sum of 5 or greater indicates a poor sleeper. In addition to a global score, the PSQI yields scores on seven composite measures that reflect sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medication, and daytime dysfunction.[13]

Big Five Inventory-II. A shortened form of the Big Five Inventory,[14] the BFI-II,[15] was administered to yield scores on five major trait domains of personality: extraversion, neuroticism, agreeableness, conscientiousness, and openness. The self-report questionnaire includes 60 questions rated on a 1 to 5 scale where a 1 equals "disagree strongly" and a 5 equals "agree strongly."

Sensation-Seeking Scale-V (SSS-V). The SSS-V is a 40-item self-report questionnaire that uses a forced-choice format to assess individual differences in optimal levels of stimulation.[16] For example, one choice might be "I would like to learn to fly an airplane" (indicating greater sensation-seeking) versus "I would not like to learn to fly an airplane" (indicating lower sensation-seeking).[17] The SSS-V can be scored as a general measure of sensation-seeking by summing all items. Four 10-item factors are also derived: Thrill and Adventure Seeking (TAS), Experience Seeking (ES), Dis-inhibition (DIS), and Boredom Susceptibility (BS).

Urgency, Premeditation, Perseverance, Sensation Seeking, and Positive Urgency Impulsivity Behavioral Scale (UPPS-P). The UPPS-P is a self-report measure of impulsivity.[18] The test consists of 59 four-point Likert scale items that are scored to yield scores on multiple facets of impulsivity:

- Negative Urgency – the tendency to give in to impulses when experiencing negative emotions;

[12] Buysse, D.J., Reynolds III, C.F., Monk, T.H., Berman, S.R., & Kupfer, D.J. (1989). "The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research." *Journal of Psychiatric Research*, 28 (2), 193-213.
[13] Details on the PSQI are available online at: http://www.sleep.pitt.edu/content.asp?id=1484&subid=2316
[14] Digman, J. M. (1990). "Personality structure: Emergence of the five-factor model." *Annual Review of Psychology*, 41 (1), 417-440.
[15] Soto, C. J., & John, O. P. (2017). "The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power." *Journal of Personality and Social Psychology*, 113, 117-143.
[16] Zuckerman, M., et al. (1964). "Development of a sensation-seeking scale." *Journal of Consulting Psychology*, 28 (6), 477. Zuckerman, M. (1996). "The psychobiological model for impulsive unsocialized sensation seeking: A comparative approach." *Neuropsychobiology*, 34, 125-129.
[17] Additional information available online at: https://scienceofbehaviorchange.org/measures/zuckerman-sensation-seeking-scale-v/
[18] Carlson, S. R., Pritchard, A. A., & Dominelli, R. M. (2013). "Externalizing behavior, the UPPS-P Impulsive Behavior scale and Reward and Punishment Sensitivity." *Personality and Individual Differences*, 54 (2), 202-207; Cyders, M. A. (2013). "Impulsivity and the Sexes: Measurement and Structural Invariance of the UPPS-P Impulsive Behavior Scale." *Assessment*, 20 (1), 86-97; Magid, V., & Colder, C. R. (2007). "The UPPS Impulsive Behavior Scale: Factor structure and associations with college drinking." *Personality and Individual Differences*, 43 (7), 1927-1937; Whiteside, S. P., & Lynam, D. R. (2001). "The Five Factor Model and impulsivity: using a structural model of personality to understand impulsivity." *Personality and Individual Differences*, 30 (4), 669-689.

- Positive Urgency – the tendency to give in to impulses when experiencing positive emotions;
- Lack of Premeditation – the tendency to act without consideration of the potential consequences of the behavior;
- Sensation-Seeking – interest in pursuing activities that are exciting and novel; and
- Lack of Perseverance – the tendency to give up in the face of boredom, fatigue and frustration.[19]

Matrix Reasoning Test. The Matrix Reasoning subtest from the Wechsler Abbreviated Scale of Intelligence (WASI-II)[20] was selected for inclusion as it is a well-validated and reliable measure of nonverbal reasoning in individuals from ages 6 to 90. On each of 30 task trials, the examinee views a visuospatial array consisting of a series of stimuli. The series is incomplete, and the participant must select from a list of four alternatives to complete it. Each item is scored as correct or incorrect. Testing ends when the participant fails three consecutive items. The raw score (total number correct) across completed trials is tabulated and translated, using a normative database, into a standardized T-score. The standardized T-score mean is 50.0 with a standard deviation of 10.0, corresponding to a fluid reasoning/IQ score of 100 or in the average range. The Matrix Reasoning task can be completed in under ten minutes for most participants. It measures fluid reasoning, visual intelligence, part-whole spatial reasoning, perceptual organization, attention to visual detail, and sequencing. This task was completed using standardized procedures with paper/pencil administration.

Trail-Making Test. The Trail-Making Test is a two-part measure of sequencing and cognitive flexibility.[21] On the first block of the task (Part A), participants must draw lines to interconnect numbers in an appropriate sequence (e.g., 1-2-3-4…26). In the second block of the test (Part B), participants alternate between numbers and letters (e.g., 1-A-2-B-3-C etc.). An automated version was implemented and required the participant to move a computer mouse to connect the numbers/letters.[22] The number of errors and completion times for each block were recorded. High numbers of errors or long completion times on Part B of the task are associated with deficient levels of executive function and cognitive flexibility.

Self-Ordered Pointing Test (SOPT). The SOPT is a measure of working memory and behavioral organization.[23] The implementation by millisecond.com uses representational drawings and the item set described in *A Compendium of Neuropsychological Tests.*[24] On each

[19] Additional information available online at: http://www1.psych.purdue.edu/~dlynam/uppspage.htm
[20] Wechsler, D. (2011). "Wechsler Abbreviated Scale of Intelligence-II." San Antonio, TX: NCS Pearson Assessments.
[21] Pearson, M. R., Murphy, E. M., Doane, A. N. (2013). "Impulsivity-like traits and risky driving behaviors among college students." *Accident Analysis & Prevention*, 53, 142-148; Petrides, M., & Milner, B. (1982). "Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man." *Neuropsychologia*, 20, 249-262; Tombaugh, T.N.T.N (2004). "Trail Making Test A and B: Normative data stratified by age and education." *Archives of Clinical Neuropsychology*, 19 (2), 203–214.
[22] See www.millisecond.com for more information.
[23] Gillett, R. (2007). "Assessment of working memory performance in self-ordered selection tests." Cortex, 43, 1047-1056. Petrides, M., & Milner, B. (1982). "Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man." *Neuropsychologia*, 20, 249-262; Ross, T. P., Hanouskova, E., Giarla, K., Calhoun, E., Tucker. M. (2007). "The reliability and validity of the self-ordered pointing task." *Archives of Clinical Neuropsychology*, 22, 449-458; Shimamura, A. P. & Jurica, P. J. (1994). "Memory interference effects and aging: Findings from a test of frontal lobe function." *Neuropsychology* 8 (3), 408-412; West, R., Ergis, A. M., Winocur, G., Saint-Cyr, J. (1998). "The contribution of impaired working memory monitoring to performance of the self-ordered pointing task in normal aging and Parkinson's disease." *Neuropsychology*, 12, 546-554.
[24] Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms and commentary (3rd Edition)*. Oxford University Press.

block of the task, participants see an array of pictured objects on the screen. On each trial, the screen changes to show the same objects but in different positions. Across trials within a given array, participants must select each object once as each screen advances without repeating any selections. That is, they must select each object only once by remembering what they have chosen on the previous screens. The object arrays vary in difficulty from 6-12 items, and the total number of errors for each difficulty level is recorded.

Multi-Source Interference Task (MSIT). The MSIT was developed to measure response control under conditions of cognitive interference.[25] On each trial, participants are presented with stimuli consisting of three digits (1, 2, 3 or one digit and two Xs as controls). One of the digits is different from the flanking digits and differs in size. Participants are asked to press a response button (1, 2, 3) as fast as possible to report which digit was different. If the selected digit is in the same button location (e.g., a 1 must be selected by pressing button 1), the trial is considered congruent. Response conflict is created when the response button (1, 2, 3) differs from the digit that must be selected (e.g., a 2 is different but must be selected by pressing button number 3), making those trials more difficult. These are referred to as "conflict/incongruent" trials. Response latencies and accuracy scores for congruent versus incongruent trials are tabulated. Relatively more errors or greater response latencies to respond under conflict/incongruent relative to non-conflict/congruent conditions are indicative of impairments in cognitive control.

Attentional Network Test-Short Form (ANT-S). The ANT-S similarly measures cognitive control under congruent versus conflict conditions[26] and has been used to index attentional control in a number of clinical populations.[27] A validated short form of the test was implemented for this assessment.[28] On each trial, participants are presented with arrows either directly above or below a fixation cross. Arrows are flanked by other arrows that point in the same (congruent) or

---

[25] Bush, G., Shin, L. M., Holmes, J., Rosen, B. R., & Vogt, B. A. (2003). "The Multi-Source Interference Task: validation study with fMRI in individual subjects." *Molecular Psychiatry*, 8, 60-70; Bush, G., & Shin, L. M. (2006). "The Multi-Source Interference Task: an fMRI task that reliably activates the cingulo-frontal-parietal cognitive/attention network." *Nature Protocols*, 1 (1), 308-13; Bush, G., Spencer, T. J., Holmes, J., Shin, L. M., Valera, E. M., Seidman, L. J., Makris, N., ... Biederman, J. (2008). "Functional magnetic resonance imaging of methylphenidate and placebo in attention-deficit/hyperactivity disorder during the multi-source interference task." *Archives of General Psychiatry*, 65 (1), 102-14; Gaetz, W., Liu, C., Zhu, H., Bloy, L., & Roberts, T. P. (2013). "Evidence for a motor gamma-band network governing response interference." *Neuroimage*, 74, 245-53; Wenzel, M., Kubiak, T., & Conner, T. S. (2013). "Positive affect and self-control: Attention to self-control demands mediates the influence of positive affect on consecutive self-control." *Cognition and Emotion*, 28 (4), 747-55.

[26] Fan, J., Gu, X., Guise, K.G., Liu, X., Fossella, J., Wang, H. & Posner, M.I. (2009). "Testing the behavioral interaction and integration of attentional networks." *Brain and Cognition*, 70, 209–220; Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). "Testing the efficiency and independence of attentional networks." *Journal Cognitive Neuroscience*, 14 (3), 340-7; Morrow, P. C. & Crum, M. R. (2004). "Antecedents of fatigue, close calls, and crashes among commercial motor-vehicle drivers." *Journal of Safety Research*, 35, 59-69; Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Pappert Lercari, L., & Posner, M.I (2004). "Development of attentional networks in childhood." *Neuropsychologia*, 42, 1029-1040; Wang, Y.F., Jing, X.J., Liu, F., Li, M.L., Long, Z-L., Yan, J.H. & Chen, H-F. (2015). "Reliable attention network scores and mutually inhibited internetwork relationships revealed by mixed design and non-orthogonal method." *Scientific Reports*, 5, 1025, doi: 10.1038/srep10251.

[27] Gooding, D. C., Braun, J. G., & Studer, J. A. (2006). "Attentional network task performance in patients with schizophrenia-spectrum disorders: Evidence of a specific deficit." *Schizophrenia Research*, 88, 1-3; López, S. G., Fuster, J. I., Reyes, M. M., Collazo, T. M., Quiñones, R. M., Berazain, A. R., Rodríguez, M. A., … Valdés-Sosa, M. (2011). "Attentional network task in schizophrenic patients and their unaffected first degree relatives: a potential endophenotype." Actas Españolas De Psiquiatría, 39, 1; Salo, R., Gabay, S., Fassbender, C., & Henik, A. (2011). "Distributed attentional deficits in chronic methamphetamine abusers: Evidence from the Attentional Network Task (ANT)." *Brain and Cognition*, 77 (3), 446-452.

[28] Weaver, B., Bédarda, M., and McAuliffe, J. (2013). "Evaluation of a 10-minute version of the Attention Network Test." *The Clinical Neuropsychologist*, 27 (8), 1281–1299.

opposite direction (incongruent). Participants have to decide whether the central arrows point right or left. Several cue conditions may alert the participants that the targets are about to come on screen and potentially where on the screen (above or below the fixation cross). Response latencies and accuracies for congruent versus incongruent trials are contrasted to yield an overall Orienting Effect (the effect of orienting participant attention toward a specific location) as well as an overall Conflict Effect (the extent to which the participant can respond efficiently under conflict conditions).

A complete list of the study measures and primary behavior variables available for data analysis can be found in Table A1 in the Appendix. The measures were administered in a uniform order across participants.

Safety Outcomes

Three independent sources were utilized to quantify participating drivers' safety performance: Motor Vehicle Records (MVR); Pre-Employment Screening Program (PSP) reports; and participant self-report.

MVR Data. Participants signed release forms authorizing ATRI staff to access their MVR for the past five years. MVR data included numbers and types of moving violations, if any, as well as DOT-reportable crash involvement. MVRs were available for each of the 94 participants. Sixty-one had no recorded MVR violations. Of the 33 participants with violations on their MVR, 20 had one violation, six had two violations, four had three violations, one had five violations, and two had six violations. There were nine individuals who had missing data regarding crash involvement. Of the remaining 76 participants, MVR-reported crashes were rare; seven participants of 76 with available data had any MVR-reported crash involvement (positive or negative). The remaining 69 individuals had no MVR-based crash involvement data available for analysis.

PSP Data. The research team secured PSP reports for 52 participants. Among these 52 individuals, numbers of DOT-reportable crashes were rare: 45 had no crashes and seven had crash involvement, although none were fatal crashes. Two were injury crashes. Other PSP indices are reported in Table 1 below. Analyzing across the various PSP metrics, drivers can be characterized as having any PSP-reported indices of poor safety (n = 25) versus having none (n = 27). The number of PSP-reported violations or crashes in the sample ranged from zero to 14.

It is important to note that drivers classified as less safe based on MVR data do not fully overlap with drivers classified as less safe based on PSP data. Of the 52 participants who had PSP data available, 25 were classified as less safe based on having any commercial vehicle-related crashes or violations. Of these 25, 11 were also classified as less safe based on having any MVR violations. When the two indices (MVR and PSP data) are combined, 11 drivers are classified as less safe on both metrics, 25 drivers are classified as less safe based on one of the metrics, and 16 drivers are classified as safe based on both metrics.

Self-Report/Participant Recollection. Participants were also asked to self-report crash and violation data. The number of participant-reported property damage only (PDO) crashes in the past five years was reported by 92 participants and the number of PDO crashes ranged from zero to three. Two cases were missing. Eighty individuals reported zero crashes, seven individuals reported one crash, four individuals reported two crashes, and one individual reported three crashes. Only two of 92 individuals reported an injury crash, and none reported fatal crashes. Inspection and moving violation data were provided by 91 participants. For inspection violations, 19 individuals reported any violations; 72 reported none. Ten participants self-reported moving violations. This self-report data was not analyzed in relation to behavioral outcomes as the data cannot be objectively verified.

Table 1 below details the safety outcomes utilized in this analysis.

Table 1. Safety Outcomes

Safety Outcome Variable	Number of Participants with Available Data	Mean (SD)	Range	Notes
Motor Vehicle Record Data				
# Motor Vehicle Violations	94	0.65 (1.22)	0 to 6	
# Motor Vehicle Crashes	76	0.11 (0.35)	0 to 2	Data available from the tradeshow samples only; of 76 individuals, 7 had experienced a crash
Pre-Employment Screening Data				
PSP Any Crashes or Violations	52	----	0 to 1	27 = none; 25 = any
PSP Crashes	52	----	0 to 2	46 = none; 5 = 1; 1 = 2
PSP Injury Crashes	52	0.04 (0.19)	0 to 1	
PSP Fatal Crashes	52	0.0 (0.0)	0 to 0	
PSP Number of Towaways	52	0.13 (.40)	0 to 3	
PSP Hazmet Releases	52	0.00 (0.00)	0 to 0	
PSP Inspection Summary	52	1.96 (2.18)	0 to 8	
PSP Vehicle Inspections	52	0.94 (1.24)	0 to 5	
PSP Hazmat Inspections	52	0.33 (0.58)	0 to 3	
PSP Violations	52	1.29 (2.36)	0 to 14	30 = 0; 17 = 1-3; 5= >3
Participant Self-Report Data				
Physical Damage Only Crashes	92	0.20 (.56)	0 to 3	
Injury Crashes	92	0.02 (.00)	0 to 1	
Fatal Crashes	92	0.00 (.00)	0 to 0	
Inspection Violations (Y/N)	91	----	----	19 yes; 72 no
Moving Violations (Y/N)	91	----	----	10 yes; 81 no

Statistical Approach

The three primary objectives of the beta test study were: a) to determine the feasibility of implementing the assessment battery in an employment setting; b) to assess whether the measures were sensitive to individual differences in driver safety in a current group of drivers across a broad age range; and c) to determine whether the measures in the battery would be sensitive to age-related performance differences.

The first objective (feasibility assessment) was assessed by examining the data for extreme or invalid values and by assessing research staff qualitative reports regarding the ease and timing of task administration. In addition, construct validation was verified by quantifying associations between measures known to show significant correlations in the literature (e.g., trait Neuroticism tends to be associated with relatively high depression levels).

For the second objective (associations with driver safety) to be achieved reliably and validly, it was necessary for safety outcomes to freely vary within the sample with adequate representations of unsafe, moderately safe, and highly safe drivers. As indicated above, however, crash events were rare among participants. Similarly, five-year MVR violations of any type were relatively infrequent in the sample overall with many zero values. PSP data were available on only a subset of participants with rare instances of commercial vehicle crashes. Thus, the sample distribution was highly skewed in favor of safe drivers. To manage this limitation in sampling, the data were examined using two approaches. The first approach utilized MVR data, as it was available on all participants, in non-parametric Spearman rank-order correlations to assess associations between absolute numbers of MVR violations and each task/questionnaire outcome. The sample was then grouped into those with no MVR violations versus any MVR violations, and outcomes were contrasted between groups using statistical probability tests. The use of these non-parametric tests circumvents the assumption that data values, e.g. crash data, are normally distributed within the sample.

A second approach was to conduct the same type of group comparisons but on the basis of any versus no PSP-related violations/crashes. Correlational analyses of the PSP data were also conducted.

The third objective (quantification of age-related variation in performance) was achieved by correlating participant age with outcome variables, including safety outcomes. For these analyses, Spearman rank-order correlations were utilized.

For all statistical analyses, two-tailed tests with an alpha level of .05 were conducted. Given the exploratory nature of the study, no corrections for multiple comparisons were implemented. Statistical interactions (e.g., between safety outcomes, age and driver characteristics) could not be reliably modeled in this study due to the small sample size.[29]

[29] Cohen, J., (1992). "A power primer." *Psychological Bulletin*, 112, 155-159.

RESULTS

Demographics

Demographic information on the study participants is presented in Table 2 below. The majority of participants were male (89%), ranging in age from 20-60, with a mean age of 44.4. Ninety-one percent self-identified as Caucasian. Years of formal education ranged from 9th-11th grade to master's level graduate education. Most participants reported having a high school or high-school-level (GED) education. Of the 94 participants, 72 reported being currently married or in a committed relationship. The average age of obtaining a CDL was 27 years and the average number of years employed as a commercial truck driver was 16.1 years, with a range from 0.67 to 39 years. As would be expected, an older current age was associated with a greater number of years employed as a commercial truck driver[30] but also with an older age of obtaining the CDL.[31]

Several employment characteristics were assessed. Drivers largely reported working for fleets that were either moderate (51 to 250 power units: 26.1%) or large (1,000+ power units: 46.6%) in size. A total of 14 drivers worked for fleets with fewer than 50 power units; five drivers worked for fleets sized between 250 and 1,000 power units. Average lengths of haul were described as local (19.2%), regional (43.6%), inter-regional (20.2%), or long-haul (17.0%).

Table 2. Study Participant Demographics

Gender	Percent
Male	89.4%
Female	10.6%
Age	
Under 25	8.5%
25-34	16.0%
35-44	17.0%
45-60	58.5%
Industry Segment	
For Hire	83.0%
Private	9.6%
Don't Know	7.4%
For-Hire Sector	
Truckload	41.9%
Specialized/ Flatbed	47.7%
Specialized/ Tanker	0.0%
Less-than-Truckload	4.7%
Other	5.8%
Employment Type	
Company Driver	91.4%
Owner-Operator with Own Authority	0.0%
Independent Contractor/ Owner-Operator Leased to Motor Carrier	8.6%

[30] r=65, p=.000
[31] r=.23, p=.026

Average Length of Haul	
Local (< 100 miles per trip)	19.2%
Regional (100 – 499 miles per trip)	43.6%
Interregional (500 – 999 miles per trip)	20.2%
Long-Haul (1,000+ miles per trip)	17.0%
Fleet Size	
≤ 5 power units	5.7%
6-15 power units	3.4%
16-50 power units	6.8%
51-250 power units	26.1%
251-500 power units	1.1%
501-1,000 power units	4.6%
> 1,000 power units	46.6%
Unknown	5.7%

Assessment Feasibility

Overall, it proved feasible to administer the full battery of tasks to current commercial truck drivers. Data values were generally reasonable and conform to expectations regarding associations between the assessed constructs. For instance, and as expected, a high Beck Depression Inventory Score was associated with poor sleep as measured by the PSQI Global Score,[32] with high levels of Big Five Inventory neuroticism,[33] and with low levels of Big Five Inventory Extraversion.[34] Negative emotion, as measured by the Big Five Inventory, correlated positively with negative urgency as measured by the UPPS-P impulsivity questionnaire.[35] Low levels of Conscientiousness, as measured by the Big Five Inventory, were associated with high levels of UPPS-P impulsivity,[36] with high scores on the AUDIT,[37] and with UPPS-P lack of perseverance.[38] Sensation-seeking, as measured by the Zuckerman SSS-V, correlated with UPPS-P total impulsivity.[39] Across cognitive measures, reaction time scores were significantly inter-correlated.[40] These associations align with the existing literature and affirm that expected relationships exist among the constructs of interest.

However, some expected associations failed to emerge as significant. For instance, Extraversion as measured by the Big Five Inventory-II was not associated with sensation-seeking traits, with high scores on the AUDIT, or with the tendency to act impulsively under conditions of elevated positive emotion (e.g., positive urgency), as might have been expected. It may be that these associations are more likely to emerge as significant in undiluted younger samples.

While the self-report questionnaire data generally adhered to expectations regarding associations among constructs, several of the cognitive task measures demonstrated questionable validity. One measure, the Trail Making Test, yielded data with considerably

32 r=.417
33 r=.465
34 r=-.439
35 r=.532
36 r=-.509
37 r=-.301
38 r=-.652
39 r=.583
40 r values were between .2 and .4

higher numbers of errors than tend to be observed in general population samples. In addition, distributions of data within several of the other cognitive measures (e.g., the ANT-S) are characterized by numerous extreme values. Given the mechanisms through which the data were collected, it is difficult to know if these limitations and challenges to feasibility are due to convenience samples, the testing environment (tradeshow event) or to limitations associated with the specific tasks. The MATS and GATS tradeshow events presented testing environments that were less than optimal for the collection of cognitive data, including background noise and some unpredictable distractions. Moreover, given that the participating drivers had other demands on their time, there was, at times, limited enthusiasm for the assessment noted by the research staff. Measures that were administered later in the battery (e.g., the ANT-S and MSIT tasks) were met with less enthusiasm than those that were administered earlier.

Table A2 in the Appendix summarizes numbers of missing cases, numbers of cases where there were concerns about the validity of the data due to lack of participant cooperation, and numbers of cases with outlier values. There were three cases excluded from all cognitive data analyses due to concerns expressed by research staff about those participants' levels of cooperation.

Driver Behavior and Safety Outcomes

MVR-Based Safety Assessment

The absolute number of recorded MVR violations was correlated with each outcome measure using Spearman Rank-Order correlations. As previously noted, this non-parametric statistical approach is relatively robust to non-normal distributions. The data were analyzed with and without outlier values, as indicated in Table A3 in the Appendix. Interpretations are not altered as a function of whether outlier cases are included in the analyses. A younger driving age was associated with more MVR violations within the five-year period.[41] An older age of obtaining the commercial driver's license was also associated with more MVR violations,[42] consistent with the observation that fewer years of employment as a commercial truck driver were also associated with more MVR violations.[43]

As indicated in Table 3 below, four variables were modestly but significantly associated with MVR violations: the Beck Depression Inventory Total Score; the Sleep Latency score of the PSQI; the Big Five measure of Extraversion; and the Big Five measure of Conscientiousness. High levels of BDI-reported depression, high sleep latency scores (reflecting an increased length of time needed to fall asleep), low levels of Extraversion and low levels of Conscientiousness were all associated with higher numbers of MVR violations.

[41] Spearman's r=-.20, p=.05
[42] Spearman's r=.26, p=.014
[43] Spearman's r= -.34, p=.001

Table 3. Correlations between Measured Driver Behaviors and Number of Past Five Years MVR Violations

	Excludes Missing and Invalid Cases			Excludes Missing and Invalid Cases plus outliers		
Task Variable	# MVR Violations r =	n =	p-value	# MVR Violations r =	n =	p-value
Beck Depression Inventory Total	.23	94	.025	.29	91	.006
Sleep Latency	.21	94	.046	.21	92	.046
BFI-II Extraversion	-.27	89	.026	-.27	89	.026
BFI-II Conscientiousness	-.25	89	.020	-.25	89	.020

The sample was grouped into those with no (n=61) versus any (n=33) MVR violations. Those without any violations are considered relatively more safe; those with any violations are considered relatively less safe. Of the 33 "less safe" drivers, two were female. Of the 61 "more safe" drivers, eight were female. While this sex distribution is not statistically significant in the current study,[44] it is notable that ATRI's 2018 Crash Predictor model found that males were involved in 20 percent more crashes than females, on average. Additionally, the Crash Predictor model found that men were more likely than females to be convicted of reckless, careless or inattentive driving, to have speeding violations, to receive citations for failures to observe traffic signals, and to be convicted of any driving-related offense.[45]

Performance data between groups were contrasted using the non-parametric Mann-Whitney U procedure. Findings are summarized in Table A4 in the Appendix, where the full sample, excluding missing cases and invalid cases, is reported. In exact parallel with the correlational analyses, significant group differences were observed for self-reported levels of depression (Beck Depression Inventory; higher levels in the less safe group), the PSQI Sleep Latency score (greater latencies in the less safe group), Big Five Extraversion (lower levels in the less safe group), and Big Five Conscientiousness (lower levels in the less safe group). When the data were analyzed excluding outliers for each variable, these differences remained significant, and no additional significant associations emerged. A summary of significant group differences is shown in Table 4 below.

Table 4. Behavioral Characteristics of MVR-based Safe versus Less Safe Drivers

Task Variable	More Safe M (SD)	More Safe Mean Rank	Less Safe M (SD)	Less Safe Mean Rank	*U* =	p-value
Beck Depression Inventory Total	5.5 (6.6)	43.42	6.20 (3.9)	55.05	757.5	.047
Sleep Latency	0.66 (.75)	43.7	0.94 (.70)	54.50	776.5	.047
BFI-II Extraversion	43.0 (6.9)	49.07	39.9 (7.9)	37.75	680.0	.047
BFI-II Conscientiousness	48.8 (6.5)	49.59	45.3 (7.3)	36.83	650.5	.025

[44] X2= 1.12, p=.29

[45] Boris, C. & Murray, D. (2018). "Predicting Truck Crash Involvement: 2018 Update." American Transportation Research Institute.

PSP-Based Safety Assessment

PSP-reported violations and crashes were summed and correlated with each outcome measure using Spearman Rank-Order correlations (Table A5 in the Appendix). As indicated in the table, higher numbers of PSP-reported violations + crashes were marginally associated with a higher body mass index[46] and with higher levels of self-reported experience seeking,[47] as reflected by the SSS-V scale. The PSQI sleep composite scores are quantified such that low scores represent better sleep quality. Higher PSP violations/crashes were counterintuitively associated with better sleep quality[48] and with marginally lower daytime sleepiness[49] as measured by the PSQI. One possible explanation for this is that the PSP violation data was generated in the past, while the sleep reports were for the present time when the driver was assessed.

Higher violations/crashes were also associated with more sensitivity to conflict on the Multi-Source Interference Task[50] and lower accuracies on the congruent trials condition of the Attentional Network Task.[51] All other associations were non-significant. Unlike MVR violations, PSP violations/crashes were not associated with participant age,[52] with the age of obtaining the CDL,[53] or with the number of years employed as a commercial driver.[54]

The data were grouped into those without any PSP-related violations[55] versus those with any PSP-related violations,[56] and the two groups were contrasted on all measures using the Mann-Whitney U test. As with the correlational approach, those with higher numbers of PSP violations + crashes had marginally higher experience-seeking and marginally higher BMIs, but were also more sensitive to conflict conditions on the Multi-Source Interference Task.[57] Those with higher numbers of PSP violations also performed with lower accuracy on congruent trials of the Attentional Network Task.[58] Beyond these findings, there were no other significant group differences.

Post-hoc assessment: Combined PSP and MVR groups

Due to the fact that different individuals were categorized as relatively less safe on the basis of MVR versus PSP data, a combined analysis based on the 52 participants with PSP data was also explored that contrasted three groups:

- those judged to be relatively safe on the basis of both indices (n=16);
- those judged to be relatively unsafe on the basis of both indices (n=11); and
- those judged to be safe on one index but not the other (n=25).

Non-parametric Kruskal-Wallis tests were utilized to contrast the three groups. The complete findings are presented in Table A6 in the Appendix. The three groups significantly differed in

---

[46] r=.27, p=.067
[47] r=.27, p=.058
[48] r=-.30, p=.032
[49] r=-.25, p=.070
[50] r=-.34, p=.018
[51] r=-.33, p=.019
[52] r=.12, NS
[53] r=-.08, NS
[54] r=-.03, NS
[55] n=27; 3 female
[56] n=25; 3 female
[57] U=157, p=.008
[58] U=204, p=.037

self-reported Conscientiousness (Big-Five II), self-reported Agreeableness (Big Five II), and self-reported Experience-Seeking (SSS-V); the "safest" group had the highest scores on Conscientiousness and Agreeableness as compared to less safe groups, and they had the lowest Experience-Seeking scores. There was a trend-level difference for the least safe drivers to be sensitive, in terms of accuracy, to conflict on the Multi-Source Interference Task. A summary of significant variables across the three groups is shown in Table 5 below.

Table 5. Exploratory Contrast of Three Groups: MVR + PSP Safe, MVR or PSP Unsafe, MVR and PSP Unsafe

Task Variable	MVR+PSP Safe	Safe on MVR or PSP but not the Other	MVR + PSP Unsafe	Group contrast Kruskal-Wallis H	Significance (Sig)
n=	16	25	11		
BFI-II Agreeableness	49.13 (5.7)	43.88 (6.1)	46.70 (4.4)	6.93	**
BFI-II Conscientiousness	52.07 (4.5)	46.71 (7.6)	44.90 (6.1)	7.90	**
SSS-V Experience Seeking	3.5 (1.6)	4.9 (1.4)	4.7 (1.6)	7.82	**
MSIT Incongruent Trial Accuracy (proportion correct)	0.87 (.20)	0.84 (.21)	0.80 (.22)	4.54	*

Values represent means and, in parentheses, standard deviations. Kruskal-Wallis H-statistics: * = p < .10; ** = p < .05; *** = p < .01.

In a further exploratory analysis, the two most extreme groups were directly contrasted (MVR/PSP "less safe", n=11 versus MVR/PSP "more safe", n=16) using the Mann-Whitney U test. This analysis revealed that the "less safe" group had:

- significantly lower scores on Conscientiousness;[59]
- marginally higher sleep latency scores, reflecting poorer sleep latency;[60]
- trend-level elevations in SSS-V dis-inhibition;[61]
- trend-level elevations in SSS-V experience seeking;[62]
- difficulty with incongruent trials of the MSIT;[63] and
- marginally greater sensitivity to conflict trials on the MSIT, indicating difficulties with cognitive control.[64]

Sensitivity of the Battery to Age-Related Performance Variations

Participants in the current study were 20 to 60 years old, with a median age of 47 years. Sixteen drivers were below the age of 30. There was a tendency for younger drivers to be characterized as relatively less safe on the basis of higher numbers of MVR violations.[65]

[59] p=.005
[60] p=.085
[61] p=.086
[62] p=.061
[63] p=.033
[64] p=.065
[65] See Table A3; this correlation is fully significant Pearson r=-.27, if parametric statistics are utilized

Spearman rank-order correlations between current driver age and all behavioral outcome variables are presented in Table A7 in the Appendix. Correlations are presented excluding missing and potentially invalid cases and, in a second step, after excluding outlier values. As indicated in Table A7, the battery is highly sensitive to age-related performance variation. Lower ages were modestly but significantly associated with higher Beck Depression scores,[66] with higher alcohol use disorder scores,[67] and with higher sleep pathology as reflected by the Pittsburgh Sleep Quality Index global score[68] and several sleep subscale scores.

With respect to personality trait scores, lower ages in this sample were associated with lower levels of Extraversion,[69] lower levels of Conscientiousness,[70] lower levels of Agreeableness,[71] and with lower levels of Openness.[72] Studies that have tracked personality development from adolescence into adulthood suggest that most changes in major personality trait levels occur relatively early, prior to the age of 30 with more modest and gradual changes thereafter.[73] Most studies agree that Neuroticism decreases while Conscientiousness and Agreeableness broadly increase over the lifespan. Beyond that general statement, findings vary somewhat according to which instrument is used to measure personality. Soto et al collected personality self-report data, using the Big-Five Inventory, from over one million internet users who ranged in age from 10 to 65.[74] They found this same pattern for Neuroticism, Conscientiousness and Agreeableness, observing the largest increase for Conscientiousness in early adulthood.

Conscientiousness may increase into adulthood due to neurodevelopmental maturation or as a result of individuals taking on increasing levels of responsibility as they assume adult roles.[75] In terms of other traits, Soto et al and others found essentially no change in Extraversion levels after the age of 15 and increases in Openness that were maximal in adolescence and young adulthood. Thus, the current study’s finding that younger age is associated with lower levels of Conscientiousness and lower levels of Agreeableness coheres with the literature. The finding of lower levels of Extraversion at younger ages does not cohere with the literature and may reflect something distinct about the truck driver sample.

Impulsivity and sensation seeking tendencies as measured by the Zuckerman Sensation-Seeking Scale-V (SSS-V) and the UPPS-P showed the expected pattern of greater impulsivity in younger individuals.[76] On the SSS-V, a lower age was associated with higher levels of self-reported disinhibition[77] and with higher levels of boredom susceptibility.[78] On the UPPS-P, lower ages were associated with higher total impulsivity scores,[79] higher levels of sensation-

---

[66] r=-.19, p=.071
[67] AUDIT: r=-.31, p=.003
[68] r=-.29, p=.005
[69] r=.36, p=.001
[70] r=.31, p=.003
[71] r=.37, p=.000
[72] r=.23, p=.036
[73] Constantinou, E., Panayiotou, G., Konstantinou, N. (2011). “Risky and aggressive driving in young adults: Personality matters.” *Accident Analysis & Prevention*, 43, 1323-1331.
[74] Soto CJ, John OP, Gosling SD, Potter J. (2011). “Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample.” *Journal of Personality and Social Psychology*, 100, 330-48.
[75] Roberts, BW and Mroczek, D. (2008). “Personality trait change in adulthood.” *Current Directions in Psychological Science*, 17 (1), 31-35.
[76] r’s = -.18 and -.33, respectively
[77] r=-.22, p=.035
[78] r=-.28, p=.006
[79] r=-.33, p=.001

seeking,[80] and higher levels of positive urgency (the tendency to act impulsively under circumstances that have the potential to bring positive rewards).[81]

The cognitive measures each yielded a similar pattern where younger ages were associated with faster reaction times.[82] This pattern was observed for both conditions of the Trail-Making test, which required participants to complete a highly practiced sequence rapidly, for the Multi-Source Interference Task (MSIT), which required the rapid processing of visually conflicting information, and for the Attentional Network Test, which also required relatively rapid processing under conflict conditions. On the ANT, in the context of generally faster reaction times, younger participants were marginally less sensitive to attentional conflict.[83] There were no significant associations between age and performance on the self-ordered pointing test, which measures working memory.

As might be expected, current age is highly correlated with years of employment in the industry.[84] Older individuals have a greater number of years of industry employment. When the number of years of industry employment is correlated with outcome variables as described above, many associations are similar to those obtained for age, although younger ages (and not years of employment) are uniquely associated with poor global sleep and with higher scores on the AUDIT. When the number of years of industry employment is correlated with study outcomes controlling for age, the only marginally significant association that remains is between a greater number of years in the industry and a lower level of MVR violations.[85]

[80] r=-.29, p=.005
[81] r=-.27, p=.008
[82] r's = .24 to .47, see Table A7
[83] r=.22, p=.062
[84] Pearson r= .68, p=.000
[85] partial r= -.21, p=.052

CONCLUSIONS

Summary of Safety Findings

The findings from this beta test provide modest support for the idea that commercial truck drivers can be distinguished in terms of MVR and PSP safety records based on behavioral indicators. Across the two types of safety records, the measures that were statistically associated with individual differences in prior violations were the Beck Depression Inventory, the Big Five measure of personality traits, some aspects of the Zuckerman Sensation-Seeking Scale, and the Pittsburgh Sleep Quality Index as well as participant age. In general, and counter to expectations, cognitive functioning was not reliably associated with histories of violations or crashes in this group of drivers, though the Multi-Source Interference Task yielded differences in relation to PSP data.

As observed in this study, low self-reported Conscientiousness has been associated in numerous reports with high numbers of workplace accidents.[86] In driving-related studies, low conscientiousness has been robustly associated with accident involvement, carelessness and a lack of behavioral control that may be indicative of a disregard for rules and regulations.[87] The construct of conscientiousness also reflects, in part, the tendency to act impulsively (low scores), and indeed, in this sample, low conscientiousness was significantly associated with higher levels of self-reported impulsivity as measured by the UPPS-P. One question that arises in evaluating the personality findings relates to whether the instruments used in the current study can reliably index population-level norms. That is, can one infer that if an individual produces a given score on the Conscientiousness scale, then that individual possesses a mid-range versus high or low value relative to others in the general population? To make such inferences, researchers would need to be able to reference obtained scores against population-level norms that take factors such as age and other demographics into account. Norm-referenced scaling for the brief form of the Big Five Inventory that was used in the current study is not yet available, so conclusions cannot be drawn regarding whether participants in the current study are, in a larger sense, relatively higher in Conscientiousness or not. The longer form of the Big Five Inventory (not selected for inclusion in this study due to time constraints) does have limited normative data available from a large sample of internet users.[88]

The observed association between low extraversion levels and higher MVR violations is less consistent with the literature. Many studies have found that high extraversion is associated with greater accident involvement due to the sensation-seeking characteristics that tend to characterize those who are relatively more extraverted.[89] However, in this sample, extraversion was not positively associated with sensation-seeking as measured by either the SSS-V or by

---

[86] Cellar, D. F., Nelson, Z. C., Yorke, C. M. (2000). "The five-factor model and driving behavior: Personality and involvement in vehicular accidents." *Psychological Reports*, 86, 454-456; Clarke, S. & Robertson, I. (2005). "A meta-analytic review of the Big Five personality factors and accident involvement in occupational and non-occupational settings." *Journal of Occupational and Organizational Psychology*, 78, 355-376; Hansen, C. P. (1988). "Personality characteristics of the accident involved employee." *Journal of Business and Psychology*, 2 (4), 346-365.
[87] Clarke, S. & Robertson, I. (2005). "A meta-analytic review of the Big Five personality factors and accident involvement in occupational and non-occupational settings." *Journal of Occupational and Organizational Psychology*, 78, 355-376.
[88] Soto CJ, John OP, Gosling SD, Potter J. (2011). "Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample." *Journal of Personality and Social Psychology*, 100, 330-48.
[89] Clarke, S. & Robertson, I. (2005). "A meta-analytic review of the Big Five personality factors and accident involvement in occupational and non-occupational settings." *Journal of Occupational and Organizational Psychology*, 78, 355-376; Lajunen, T. (2001). "Personality and accident liability: Are extraversion, neuroticism and psychoticism related to traffic and occupational fatalities?" *Personality and Individual Differences*, 31, 1365-1373.

the UPPS-P.[90] To some extent, relatively low rather than high levels of extraversion might be considered beneficial to success as a truck driver given the solitary nature of the work.[91] To the extent that low extraversion is a risk factor for higher numbers of driving-related violations, it may be due to the fact that low extraversion can confer low levels of social or occupational engagement, which may be disruptive to motivation and attentional focus. Low levels of positive affect, as seen in depressive conditions, may also contribute to this association and would align with the observations that extraversion was negatively associated with Beck Depression scores and that relatively higher levels of self-reported depression symptoms were associated with greater numbers of MVR violations. Relatively higher depression levels, in addition to impacting motivation, may lead to decreased arousal, slowed reaction times, fatigue and poor attention/focus, all of which could contribute to careless driving behavior.[92] A critical caveat in evaluating the current data is that despite yielding relatively higher scores on the BDI, participants in the current study who had greater numbers of prior MVR violations were not, on average, clinically depressed. Their depression scores were solidly in the normative range with only three individuals in the entire sample scoring above that level.

A longer self-reported sleep latency (reflecting a longer period of time to fall asleep), the facet of the PSQI that was most strongly correlated with the global sleep quality score, was modestly associated with greater numbers of MVR violations. In the current sample, this is an isolated finding. Other measures such as nighttime sleep duration, sleep efficiency, daytime sleepiness and use of sleep medications did not show significant associations with MVR violations. Moreover, it is not necessarily the case that individuals with greater sleep latencies have other sleep disturbances.[93] Thus, the correlation of sleep latency with MVR violations could be due to other lifestyle factors. Causal modeling cannot be statistically attempted in this study due to the small sample size.

Nonetheless, consistent with this finding, various indices reflective of poor sleep and driver fatigue have been associated with increased risks for unsafe driving behavior and elevated crash risk in numerous studies.[94] Many models describing associations between poor sleep and driver errors suggest that attentional processes are adversely impacted by the fatigue state.[95] Given that attentional difficulties have been associated in other studies with adverse

---

[90] Appendix II, Tables C and E.
[91] Knipling, R. R., Burks, S.V., Starner, K.M. et al. (2011). "Driver Selection Tests and Measurement." CTBSSP Synthesis of Safety Practice No.21. Transportation Research Board of the National Academies, Washington, D.C.
[92] Bunn, T.L., Slaboba, S., Struttmann, T.W., Browning, S.R., (2005). "Sleepiness, fatigue and distraction/inattention as factors for fatal versus nonfatal commercial motor vehicle driver injuries." *Accident Analysis and Prevention*, 37, 862–869; Dinges DF. (1995). "An overview of sleepiness and accidents." *Journal of Sleep Research*, 4 (S2), 4–14. Jennings, J. R.; Monk, T. H.; der Molen, M. W. (2003). "Sleep deprivation influences some but not all processes of supervisory attention." *Psychological Science*, 14 (5), 473–479; Jerome, L., Segal, A., Habinski, L. (2006). "What we know about ADHD and driving risk: A literature review, meta-analysis and critique." *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 15 (3), 105-125.
[93]Appendix II, Table B.
[94] Braver, E.R., Preusser, C.W.Preusser, D.F., Baum, H.M., Beilock, R., Ulmer, R., (1992). "Long hours and fatigue: a survey of tractor-trailer drivers." *Journal of Public Health Policy*, 13 (3), 341–366; Knipling, R.R. & Wang, J.S. (1994). "Crashes and fatalities related to driver drowsiness/fatigue." NHTSA Research Note, November 1994. McCartt, A. T., Rohrbaugh, J. W., Hammer, M. C., & Fuller, S. Z. (2000). "Factors associated with falling asleep at the wheel among long-distance truck drivers." *Accident Analysis & Prevention*, 32, 493-519; Moore-Ede, M., Heitmann, A., Guttkuhn, R., Trutschel, U., AGuirre, A., Croke, D. (2004). "Circadian alertness simulator for fatigue risk assessment in transportation: Application to reduce frequency and severity of truck accidents." *Aviation, Space, and Environmental Medicine*, 75 (Supplement 1), A107-A118; Morrow, P. C. & Crum, M. R. (2004). "Antecedents of fatigue, close calls, and crashes among commercial motor-vehicle drivers." *Journal of Safety Research*, 35, 59-69.
[95] Jennings, J. R.; Monk, T. H.; der Molen, M. W. (2003). "Sleep deprivation influences some but not all processes of supervisory attention." *Psychological Science*, 14 (5), 473-479.

driving outcomes,[96] it is surprising and somewhat unexpected that the measures of executive function and attentional control that were administered in this study failed to show any significant association with histories of MVR violations and crashes. Commercial driving is a skill that involves considerable multi-tasking ability, inhibitory control,[97] and directed attention.[98] On the one hand, it could be that the tasks used in the current study lacked sensitivity.[99] Conversely, it may be that the types of violations demonstrated by study participants did not reflect the types of incidents that would be associated with individual differences in executive function. For instance, failures of inhibitory control may be associated with specific types of safety violations (e.g., failure to brake, hard braking events, speeding), which could not be statistically differentiated among the violations made by this study's participants.

There was some suggestion of an association between PSP-recorded violations and diminished levels of cognitive control as measured by the Multi-Source Interference Task; this task also distinguished extreme groups of individuals characterized by safe versus less safe driving histories on the basis of combined MVR and PSP data, though sample sizes were small. Given that PSP-recorded violations reflect incidents that have occurred in commercial driving contexts, they might be considered to be more serious in nature. Thus, this association merits further investigation.

The literature is consistent in suggesting associations between elevated sensation-seeking and crash risk, as well as similar associations between elevated levels of impulsivity and indices of poor safety.[100] In the current study, these behaviors, which were measured using well-validated tests, surprisingly did not distinguish individuals with relatively better versus worse safety records. Only one facet of sensation-seeking (Experience-seeking, as measured by the SSS-V) was shown to be distinct in individuals with relatively high versus lower PSP-recorded violations/crashes, though this association was statistically tenuous and should not be over-interpreted, particularly given the limited sample size. Experience-seeking reflects the tendency to pursue or explore novel environments. Other studies that have used the SSS-V (the most commonly employed measure of sensation-seeking in the driving literature[101]) have found

[96] Barkley, R.A. (2004). "Driving impairments in teens and adults with attention-deficit/hyperactivity disorder." *Psychiatric Clinics of North America*, 27 (2), 233-60; Barkley, RA, Murphy, KR, Dupaul, GI, Bush, T. (2002). "Driving in young adults with attention deficit hyperactivity disorder: knowledge, performance, adverse outcomes, and the role of executive functioning." *Journal of the International Neuropsychological Society*, 8 (5), 655-72.
[97] O'Brien, F. & Gormley, M. (2013). "The contribution of inhibitory deficits to dangerous driving among young people." *Accident Analysis & Prevention*, 51, 238-242.
[98] Roca, J., Crundall, D., Moreno-Ríos, S., Castro, C., Lupiáñez, J. (2013). "The influence of differences in the functioning of the neurocognitive attentional networks on drivers' performance." *Accident Analysis & Prevention*, 50, 1193-1206.
[99] Barrash, J., Stillman, A., Anderson, S. W., Uc, E. Y., Dawson, J. D., Rizzo, M. (2010). "Prediction of driving ability with neuropsychological tests: Demographic adjustments diminish accuracy." *Journal of the International Neuropsychological Society*, 16 (4), 679.
[100] Dahlen, E. R., Martin, R. C., Ragan, K., Kuhlman, M. M. (2005). "Driving anger, sensation seeking, impulsiveness, and boredom proneness in the prediction of unsafe driving." *Accident Analysis & Prevention*, 37 (2), 341-348; Dahlen, E. R. & White, R. P. (2006). "The Big Five factors, sensation seeking, and driving anger in the prediction of unsafe driving." *Personality and Individual Differences*, 41 (5), 903-915; Iversen, H., Rundmo, T., (2002). "Personality, risky driving and accident involvement among Norwegian drivers." *Personality and Individual Differences*, 33 (8), 1251-1263; Jonah, B.A., (1997). "Sensation seeking and risky driving: a review and synthesis of the literature." *Accident Analysis & Prevention* 29, 651–665; Knipling, R. R., Burks, S.V., Starner, K.M. et al. (2011). "Driver Selection Tests and Measurement." CTBSSP Synthesis of Safety Practice No.21. Transportation Research Board of the National Academies, Washington, D.C.; Pearson, M. R., Murphy, E. M., Doane, A. N. (2013). "Impulsivity-like traits and risky driving behaviors among college students." *Accident Analysis & Prevention*, 53, 142-148.
[101] Knipling, R. R., Burks, S.V., Starner, K.M. et al. (2011). "Driver Selection Tests and Measurement." CTBSSP Synthesis of Safety Practice No.21. Transportation Research Board of the National Academies, Washington, D.C.;

robust associations between high levels of sensation-seeking and risk-taking behaviors such as speeding, aggressive driving, and driving under the influence of alcohol as well as, more specifically, between thrill-seeking and disinhibition and self-reported driving offenses in non-commercial drivers.[102] Several studies have focused on young adult drivers.[103]

While the variation in findings based on whether MVR data or PSP data were utilized raises questions about the reliability of observed associations, it should be noted that the PSP data are limited given that only 52 of 94 individuals provided access to their PSP report. Thus, the two sets of findings are not directly comparable. Additional limitations, described below, are also likely to have impacted the detection of significant effects.

Sensitivity to Age

By design, the current study does not directly address the utility of this battery in the prediction of safety outcomes in young drivers, which is the ultimate objective of research. Of the current participants, only 16 drivers were below the age of 30. Thus, it is not known whether a different set of behavioral variables might emerge as significant from a similar larger study of young drivers. The fact that the battery is sensitive to age-related variations in performance is encouraging.[104] Moreover, many of the behavioral domains that showed strong age-related associations in this study are those that have been robustly associated with driving errors, lapses, and violations in studies of young drivers (e.g., impulsivity; problematic alcohol use; poor sleep quality). Given that the main effects of age as well as effects of safety histories were detected in this beta test, a logical next step would be to consider potential interactions between the two, although the current study was not adequately powered to quantify such interactions given the sampling bias in favor of relatively older drivers. Notably, of the 31 individuals with any history of MVR violations, only six were below the age of 30. Only one individual below the age of 30 had any PSP-recorded violations or crashes.

Limitations and Qualifications

While the non-cognitive aspects of the battery proved highly feasible to administer on a large scale to current drivers, the testing environments (e.g., tradeshow events) were not optimized for cognitive testing. The cognitive tests were characterized in some cases by high numbers of errors and by extreme values that are difficult to interpret. It is unclear whether they are due to characteristics of the sample, a battery that was simply too lengthy, or sub-optimal characteristics of the testing environment. Therefore, the lack of findings for measures in this domain may not generalize to other samples or settings. Undoubtedly, the testing environment should be better controlled if cognitive measures were to be utilized as part of routine employee selection.

Additionally, drivers selected for this beta test were currently employed commercial drivers and, in theory, those drivers with egregious safety records are not able to stay employed in the

Pearson, M. R., Murphy, E. M., Doane, A. N. (2013). "Impulsivity-like traits and risky driving behaviors among college students." *Accident Analysis & Prevention*, 53, 142-148.

[102] Costa, PT, McCrae, RR, Löckenhoff, CE. (2019). "Personality across the lifespan." *Annual Review of Psychology*, 70 (1), 423-448; Jonah, B.A., (1997). "Sensation seeking and risky driving: a review and synthesis of the literature." *Accident Analysis & Prevention*, 29, 651-665.

[103] Costa, PT, McCrae, RR, Löckenhoff, CE. (2019). "Personality across the lifespan." *Annual Review of Psychology*, 70 (1), 423-448; Pearson, M. R., Murphy, E. M., Doane, A. N. (2013). "Impulsivity-like traits and risky driving behaviors among college students." *Accident Analysis & Prevention*, 53, 142-148.

[104] As indicated in Table A7.

industry. As a result, instances of unsafe behavior were not strongly represented in the sample, and the violations that were recorded could be construed in some cases as being relatively minimal in their potential for occupational significance.

This lack of representation creates challenges in the interpretation of the data as related to the question of whether inadequate safety profiles can be reliably detected retrospectively or predicted over time. Given that the ultimate goal of pre-employment screening would be to exclude individuals predicted to demonstrate high levels of violations, crashes, aggression or other risk-taking driving behaviors (or conversely to select those who are likely to be maximally reliable), any algorithms derived to enable such prediction must be based on large numbers of individuals with both types of records. That is, a full continuum of safety performance must be evident in the derivation sample. This was not the case for the current sample where crash incidents were low, and less than half of the participants had any MVR- or PSP-reported safety violations. Accordingly, the statistical power to detect characteristics associated with inadequate safety records, and to differentiate among more versus less serious types of violations, was low. Even when violations were observed, they tended to be few in number for individual participants (e.g., one recorded violation). The violations that were evident were sufficiently minimal in number, and as such they could not be statistically distinguished in a reliable fashion (e.g., moving violations versus vehicle inspection-related violations) in terms of their associations with driver characteristics.

In conclusion, in this small sample, the truck driving safety profiles of current truck drivers, as measured by motor vehicle records and pre-employment screening data, can be distinguished based on personality traits, physiological characteristics, and aspects of mental health. With expanded statistical validation, this methodology may be successfully applied to the assessment and selection of new entrants into the industry's workforce, including younger drivers, to assist in the identification of those who are most likely to drive safely. This is of particular importance as the industry evaluates the potential for younger drivers to mitigate the driver shortage.

<u>Next Steps</u>

In light of the pattern of observed findings and the associated limitations, several priorities have been identified for ATRI's expanded Phase II research in order to determine whether young individuals with safe driving behaviors can be distinguished from those with less safe driving behavior.

Increase the Sample of Young Drivers. ATRI's Phase II assessment will focus on a larger sample of young drivers by recruiting individuals in the 18- to 25-year old age range from commercial driving schools. By administering Phase II testing under the same conditions as the beta test, the two datasets can be combined in order to test for possible interactions between age, experience and safety outcomes.

Expand the Range of Participant Driver Histories. The driver recruitment process will pursue a broader range of driver histories among participants by utilizing multiple distinct types and locations of commercial driving schools. Ideally, the expanded sample will include sufficiently large numbers of individuals representing different gender, race, education and driving experience backgrounds as well as more individuals with past safety violations in order to analyze responses from comparatively less safe versus more safe drivers.

Shorten the Testing Battery. Although it is not clear whether a distinct pattern of findings might emerge among a group of younger participants, the testing battery can be truncated based on the current findings. Measures to be eliminated include:

- BMI assessment – more likely to be predictive of health risks and threats to safety such as poor sleep in older and more seasoned drivers versus newly hired individuals;
- UPPS-P – though highly sensitive to age, was not predictive of driver safety;
- Trail-Making Test – due to significant questions about validity in the context of automated administration; and
- Attentional Network Test – the ANT is redundant in theory with the Multi-Source Interference Task and was not well received by the drivers.

Executive functions will still be assessed in Phase II given data suggesting that younger people are still developing these skills into their mid-20s. In light of their utility in the beta test, measures of mental health, sleep quality, substance abuse liability, personality traits, and sensation-seeking all will be retained, as will MVR data.

Given the importance of the findings in contributing to debates regarding national policies related to truck driver safety and public health, any subsequent study that enrolls young drivers as the sample of interest should be adequately powered. A sample size of 300 individuals, distributed according to safety profiles, is being targeted in order to yield reliable and adequately powered findings for the retained assessments. ATRI's Phase II will commence in the fourth quarter of 2021.

APPENDIX

Table A1. List of Measures and Analyzed Variables

Task	Variables
Demographics Questionnaire	Age; Gender; Handedness; Level of Education; Race/Ethnicity; Language Proficiency; Relationship status; Household Configuration; Household Income; Employment Status; Industry Sector; Fleet Size; Vehicle Configuration; Age CDL; Haul Length; Annual Miles Driven
Beck Depression Inventory	Total Depression Score
Alcohol Use Disorders Identification Test (AUDIT)	Total AUDIT score
Pittsburgh Sleep Quality Index	Global Sleep Quality Score
	Subjective Sleep Quality
	Sleep Latency
	Sleep Duration
	Habitual Sleep Efficiency
	Sleep Disturbances
	Use of Sleep Medication
	Daytime Dysfunction
Body Mass Index	Height (inches)
	Weight (pounds)
	BMI {(pounds/inches) * conversion factor}
Big Five Inventory II	Extraversion
	Neuroticism
	Agreeableness
	Conscientiousness
	Openness
Zuckerman Sensation-Seeking Scale V (SSS-V)	Total Sensation-Seeking Score
	Thrill and Adventure Seeking
	Experience Seeking
	Disinhibition
	Boredom Susceptibility
UPPS-P	Positive Urgency
	Negative Urgency
	Lack of Premeditation
	Lack of Perseverance
	Sensation Seeking
	UPPS-P Total Score
Matrix Reasoning Test	Total Matrix Reasoning T-Score
Trail-Making Test	Trails A Errors
	Trails A Completion Time
	Trails B Errors

Task	Variables
	Trails B Completion Time
Self-Ordered Pointing Task	Errors 6-item set
	Errors 8-item set
	Errors 10-item set
	Errors 12-item set
	Total Errors
Multi-Source Interference Task	Congruent Trial Accuracy
	Congruent Trial Reaction Time
	Incongruent Trial Accuracy
	Incongruent Trial Reaction Time
	(Congruent – Incongruent) Accuracy
	(Congruent – Incongruent) Reaction Time
Attentional Network Test-Short	Orienting Effect
	Conflict Effect
	Congruent Trial Accuracy
	Congruent Trial Reaction Time
	Incongruent Trial Accuracy
	Incongruent Trial Reaction Time

Table A2. Data Quality for Total Sample of n=94

	a	b	c
Measure	**Missing Data n=**	**Data Potentially Invalid due to Participant Non-Compliance n=**	**Outliers n=**
Demographics Questionnaire	0	0	n/a
Beck Depression Inventory	0	0	3
Alcohol Use Disorders Identification Test (AUDIT)	1	0	2
Pittsburgh Sleep Quality Index Global Score	1	0	3
Pittsburgh Sleep Quality Index Composite Scores	0	0	0 to 10; (highest for sleep medication and habitual sleep quality)
Body Mass Index	0	0	2
Big Five Inventory II	5	0	2 (openness)
Zuckerman Sensation-Seeking Scale V (SSS-V)	0	0	3 (2 BS; 1 DIS)
UPPS-P	0	0	0
Matrix Reasoning Test	0	0	1
Trail-Making Test	0	3	4
Self-Ordered Pointing Task	0	3	3
Multi-Source Interference Task	1	3	13
Attentional Network Test-Short	1	4	12

Table A3. Correlations between Measured Driver Behaviors and Number of Past Five Years MVR Violations

	Excludes Missing and Invalid Cases				Excludes Missing and Invalid Cases plus Outliers			
Task Variable	#MVR Violations r=	n=	p-value	Sig	#MVR Violations r=	n=	p-value	Sig
Participant Age	-.20	94	.050	*				
Beck Depression Inventory Total	.23	94	.025	**	.29	91	.006	***
Body Mass Index (BMI)	.07	94	.524		.02	92	.857	
AUDIT Total Score	.07	93	.480		.11	91	.319	
PSQI Global Score	.16	93	.115		.18	90	.109	
Subjective Sleep Quality	.15	94	.159		.15	94	.159	
Sleep Latency	.21	94	.046	**	.21	92	.046	**
Sleep Duration	-.00	94	.986		.02	93	.881	
Habitual Sleep Efficiency	.11	94	.301		-.17	84	.129	
Sleep Disturbances	-.04	94	.715		-.04	94	.715	
Use of Sleep Medication	-.01	94	.913		.01	88	.937	
Daytime Dysfunction	.14	94	.181		.14	94	.181	
Matrix Reasoning T-Score	.09	94	.385		.11	93	.312	
BFI-II Extraversion	-.27	89	.026	**	-.27	89	.026	**
BFI-II Neuroticism	-.07	89	.501		-.07	89	.501	
BFI-II Agreeableness	-.14	89	.206		-.14	89	.206	
BFI-II Conscientiousness	-.25	89	.020	**	-.25	89	.020	**
BFI-II Openness	-.05	89	.645		.00	87	.994	
UPPS-P Total Score	.06	94	.552		.06	94	.552	
UPPS-P Positive Urgency	-.01	94	.924		-.01	94	.924	
UPPS-P Negative Urgency	.03	94	.774		.03	94	.774	
UPPS-P Lack of Premeditation	-.01	94	.945		-.01	94	.945	
UPPS-P Lack of Perseverance	.14	94	.175		.14	94	.175	
UPPS-P Sensation Seeking	.08	94	.443		.08	94	.443	
SSS-V Total Score	.10	94	.349		.10	94	.349	
SSS-V Thrill Seeking	.04	94	.677		.04	94	.677	
SSS-V Experience Seeking	.10	94	.347		.10	94	.347	
SSS-V Disinhibition	.13	94	.226		.13	93	.229	
SSS-V Boredom Susceptibility	.06	94	.537		.08	92	.450	

Self-Ordered Pointing Task (SOP)								
SOP Total Errors Six Item Trials	-.04	91	.735		-.10	88	.339	
SOP Total Errors Eight Item Trials	.04	91	.739		-.02	88	.827	
SOP Total Errors Ten Item Trials	.13	91	.221		.08	88	.461	
SOP Total Errors Twelve Item trials	.18	91	.098	*	.13	88	.233	
SOP Total Errors (All Trials)	.10	91	.338		.05	88	.644	
Trail-Making Test (TMT)								
TMT Part A # Errors	-.06	91	.573		-.10	87	.382	
TMT Part A Completion Time	-.11	91	.282		-.15	87	.163	
TMT Part B # Errors	-.05	91	.635		-.04	87	.710	
TMT Part B Completion Time	-.09	91	.422		-.07	87	.532	
Multi-Source Interference Task (MSIT)								
MSIT Congruent Trial Accuracy (proportion correct)	.03	90	.759		.02	77	.892	
MSIT Incongruent Trial Accuracy (proportion correct)	-.09	90	.411		-.14	77	.220	
MSIT Congruent Trial Reaction Time (ms)	-.19	90	.077	*	-.19	77	.100	
MSIT Incongruent Trial Reaction Time (ms)	-.15	90	.165		-.14	77	.223	
Attentional Network Task (ANT)								
ANT Congruent Trial Accuracy % Correct	-.09	89	.355		-.13	77	.277	
ANT Incongruent Trial Accuracy % Correct	-.03	89	.802		-.03	77	.803	
ANT Congruent Trial Reaction Time (ms)	-.16	89	.146		-.09	77	.419	
ANT Incongruent Trial Reaction Time (ms0	-.15	89	.157		-.13	77	.282	
ANT Orienting Effect	.01	89	.954		.03	77	.828	
ANT Conflict Effect	-.04	89	.701		-.02	77	.891	

r-values represent Spearman Rank-Order Correlations; *=p<.10; **=p,.05; ***=p<.01

Table A4. Behavioral Characteristics of MVR-based Safe versus Less Safe Drivers

Task Variable	More Safe M (SD)	More Safe Mean Rank	Less Safe M (SD)	Less Safe Mean Rank	*U*=	p-value	Sig
Participant Age	45.5 (12.2)	50.43	42.2 (11.4)	42.09	828.0	.158	
Beck Depression Inventory Total	5.5 (6.6)	43.42	6.2 (3.9)	55.05	757.50	.047	**
Body Mass Index (BMI)	31.6 (5.7)	45.96	33.1 (8.3)	50.35	912.50	.460	
AUDIT Total Score	2.9 (3.5)	45.65	3.3 (3.3)	49.45	909.0	.513	
PSQI Global Score	3.8 (2.5)	43.74	4.4 (2.1)	52.92	794.5	.112	
Subjective Sleep Quality	0.59 (.56)	45.3	0.73 (.52)	51.6	872.0	.244	
Sleep Latency	0.66 (.75)	43.7	0.94 (.70)	54.5	776.5	.047	**
Sleep Duration	0.33 (.57)	47.9	0.27 (.45)	46.7	979.5	.824	
Habitual Sleep Efficiency	0.41 (.90)	45.9	0.64 (1.11)	50.4	909.5	.331	
Sleep Disturbances	1.21 (.61)	48.0	1.15 (.51)	46.5	974.5	.801	
Use of Sleep Medication	0.21 (.71)	47.7	0.09 (.29)	47.1	993.0	.749	
Daytime Dysfunction	0.38 (.55)	44.6	0.58 (.61)	52.8	830.5	.124	
Matrix Reasoning T-Score	47.7 (11.0)	45.78	49.3 (9.7)	50.68	901.50	.405	
BFI-II Extraversion	43.0 (6.9)	49.07	39.9 (7.9)	37.75	680.0	.047	**
BFI-II Neuroticism	29.2 (9.1)	46.25	28.2 (9.1)	42.78	841.0	.547	
BFI-II Agreeableness	46.3 (6.6)	47.52	44.1 (6.5)	40.52	768.5	.221	
BFI-II Conscientiousness	48.8 (6.5)	49.59	45.3 (7.3)	36.83	650.5	.025	**
BFI-II Openness	43.7 (5.4)	45.95	42.8 (6.7)	43.31	858.0	.647	
UPPS-P Total Score	112.7 (23.9)	46.08	115.6 (19.2)	50.12	920.0	.496	
UPPS-P Positive Urgency	23.8 (8.3)	47.80	23.6 (7.2)	46.94	988.0	.885	
UPPS-P Negative Urgency	22.6 (7.1)	46.73	22.8 (5.5)	48.92	959.5	.712	
UPPS-P Lack of Premeditation	19.9 (4.9)	47.46	19.9 (4.6)	47.58	1004.0	.986	
UPPS-P Lack of Perseverance	16.5 (3.3)	45.10	17.4 (3.5)	51.94	860.0	.246	
UPPS-P Sensation Seeking	30.0 (7.9)	45.26	31.9 (8.1)	51.64	870.0	.281	
SSS-V Total Score	14.0 (5.7)	45.11	15.9 (6.6)	51.92	860.5	.249	
SSS-V Thrill Seeking	5.3 (2.6)	46.31	5.7 (3.1)	49.70	934.0	.566	

SSS-V Experience Seeking	4.1 (1.7)	44.98	4.5 (1.4)	52.17	852.5	.217	
SSS-V Disinhibition	2.6 (2.2)	44.56	3.3 (2.4)	52.94	827.0	.152	
SSS-V Boredom Susceptibility	2.1 (1.7)	45.96	2.4 (2.0)	50.35	912.5	.451	
Self-Ordered Pointing Task (SOP)							
SOP Total Errors Six Item Trials	1.8 (2.2)	47.14	2.2 (3.6)	43.91	877.0	.569	
SOP Total Errors Eight Item Trials	2.4 (2.9)	45.04	3.3 (4.8)	47.77	887.5	.632	
SOP Total Errors Ten Item Trials	3.4 (3.7)	43.81	4.8 (6.1)	50.05	814.5	.279	
SOP Total Errors Twelve Item trials	4.6 (4.5)	43.12	6.5 (7.4)	51.31	774.0	.156	
SOP Total Errors (All Trials)	12.3 (12.3)	44.26	16.7 (21.5)	49.20	841.5	.396	
Trail-Making Test (TMT)							
TMT Part A # Errors	1.07 (1.23)	46.65	1.03 (1.24)	44.86	919.5	.748	
TMT Part A Completion Time	54695.4 (13621.6)	47.74	53502.9 (17181.9)	42.94	856.0	.409	
TMT Part B # Errors	2.12 (2.24)	46.59	2.12 (2.27)	44.97	923.0	.777	
TMT Part B Completion Time	76725.5 (24539.4)	47.29	74696.0 (27944.3)	43.73	882.0	.541	
Multi-Source Interference Task (MSIT)							
MSIT Congruent Trial Accuracy (proportion correct)	.86 (.20)	44.8	.86 (.22)	46.8	888.0	.738	
MSIT Incongruent Trial Accuracy (proportion correct)	.83 (.22)	47.3	.83 (.22)	42.2	823.0	.378	
MSIT Congruent Trial Reaction Time (ms)	1004.0 (164.1)	48.8	943.4 (182.0)	39.5	736.0	.107	
MSIT Incongruent Trial Reaction Time (ms)	1050.3 (167.8)	48.2	989.1 (187.3)	40.1	770.0	.185	
Attentional Network Task (ANT)							
ANT Congruent Trial Accuracy % Correct	93.95 (20.5)	46.3	98.18 (2.8)	42.69	838.0	.460	

ANT Incongruent Trial Accuracy % Correct	75.6 (38.0)	44.8	72.8 (38.1)	45.31	902.0	.933	
ANT Congruent Trial Reaction Time (ms)	586.4 (106.5)	48.2	544.1 (69.0)	39.3	731.0	.123	
ANT Incongruent Trial Reaction Time (ms)	751.2 (224.1)	46.3	691.9 (207.7)	38.2	682.0	.154	
ANT Orienting Effect	11.6 (44.6)	44.9	11.4 (40.0)	45.1	907.5	.976	
ANT Conflict Effect	165.1 (173.7)	44.5	149.5 (172.7)	41.7	785.0	.623	

Values represent means +/- one standard deviation; two-tailed exact significance is reported: *=p<.10; **= p<.05; *** = p<.01. All analyses exclude missing cases and potentially invalid cases (see Table A2).

Table A5. Associations between PSP-recorded Violations + Crashes and Behavioral Outcomes

	Excludes Missing and Invalid Cases			
Task Variable	**#PSP Violations/Crashes r=**	**n=**	**p-value**	**Sig**
Participant Age	.12	52	.404	
MVR Violations	.07	52	.622	
Beck Depression Inventory Total	-.02	52	.876	
Body Mass Index (BMI)	.27	52	.067	*
AUDIT Total Score	-.19	52	.183	
PSQI Global Score	-.22	52	.112	
Subjective Sleep Quality	-3.0	52	.032	**
Sleep Latency	.06	52	.673	
Sleep Duration	-.18	52	.215	
Habitual Sleep Efficiency	.04	52	.755	
Sleep Disturbances	-.15	52	.280	
Use of Sleep Medication	-.11	52	.450	
Daytime Dysfunction	-.25	52	.070	*
Matrix Reasoning T-Score	-.14	52	.326	
BFI-II Extraversion	-.15	49	.302	
BFI-II Neuroticism	.05	49	.748	
BFI-II Agreeableness	-.09	49	.527	
BFI-II Conscientiousness	-.21	49	.150	
BFI-II Openness	.09	49	.537	
UPPS-P Total Score	-.05	52	.725	
UPPS-P Positive Urgency	-.08	52	.598	
UPPS-P Negative Urgency	-.15	52	.292	
UPPS-P Lack of Premeditation	.91	52	.175	
UPPS-P Lack of Perseverance	.06	52	.751	
UPPS-P Sensation Seeking	-.05	52	.714	
SSS-V Total Score	.05	52	.708	
SSS-V Thrill Seeking	-.08	52	.599	
SSS-V Experience Seeking	.27	52	.058	*
SSS-V Disinhibition	.12	52	.409	
SSS-V Boredom Susceptibility	-.08	52	.578	

Self-Ordered Pointing Task (SOP)				
SOP Total Errors Six Item Trials	-.18	49	.230	
SOP Total Errors Eight Item Trials	.20	49	.162	
SOP Total Errors Ten Item Trials	.05	49	.726	
SOP Total Errors Twelve Item trials	.08	49	.587	
SOP Total Errors (All Trials)	.10	49	.518	
Trail-Making Test (TMT)				
TMT Part A # Errors	.05	49	.746	
TMT Part A Completion Time	.08	49	.580	
TMT Part B # Errors	-.11	49	.462	
TMT Part B Completion Time	.03	49	.815	
Multi-Source Interference Task (MSIT)				
MSIT Congruent Trial Accuracy (proportion correct)	-.02	48	.879	
MSIT Incongruent Trial Accuracy (proportion correct)	.10	48	.499	
MSIT Congruent Trial Reaction Time (ms)	-.10	48	.481	
MSIT Incongruent Trial Reaction Time (ms)	-.19	48	.189	
MSIT Conflict (Reaction Time)	-.34	48	.018	**
MSIT Conflict (Accuracy)	.09	48	.539	
Attentional Network Task (ANT)				
ANT Congruent Trial Accuracy % Correct	-.33	49	.019	**
ANT Incongruent Trial Accuracy % Correct	.11	49	.445	
ANT Congruent Trial Reaction Time (ms)	.19	49	.202	
ANT Incongruent Trial Reaction Time (ms)	.05	49	.715	
ANT Orienting Effect	-.02	49	.901	
ANT Conflict Effect	-.06	49	.685	

Table A6. Exploratory contrast of Three Groups: MVR + PSP Safe, MVR or PSP Unsafe, MVR and PSP Unsafe

Task Variable	MVR+PSP Safe	Safe on MVR or PSP but not the Other	MVR + PSP Unsafe	Group Contrast Kruskal-Wallis H	Sig
n=	16	25	11		
Participant Age	47.81 (10.2)	45.60 (11.3)	44.18 (10.9)	0.47	
Beck Depression Inventory Total	5.69 (5.39)	4.68 (4.07)	6.91 (4.39)	2.15	
Body Mass Index (BMI)	30.90 (4.75)	31.10 (4.71)	35.23 (10.86)	1.83	
AUDIT Total Score	2.88 (2.85)	2.16 (2.17)	1.82 (1.60)	0.61	
PSQI Global Score	3.9 (1.9)	3.7 (2.4)	4.4 (2.5)	0.54	
Subjective Sleep Quality	0.75 (.45)	0.56 (.58)	0.64 (.51)	1.59	
Sleep Latency	0.56 (.73)	0.72 (.74)	1.09 (.83)	3.39	
Sleep Duration	0.50 (.52)	0.32 (.48)	0.18 (.51)	3.00	
Habitual Sleep Efficiency	0.13 (.34)	0.60 (.96)	0.55 (1.21)	3.09	
Sleep Disturbances	1.25 (.58)	1.00 (.58)	1.27 (.65)	2.16	
Use of Sleep Medication	0.19 (.75)	0.12 (.33)	0.09 (.30)	0.29	
Daytime Dysfunction	0.56 (.63)	0.40 (.65)	0.55 (.52)	1.44	
Matrix Reasoning T-Score	49.50 (9.2)	50.60 (7.1)	47.09 (10.9)	0.37	
BFI-II Extraversion	44.33 (5.2)	42.33 (6.1)	39.80 (9.8)	2.51	
BFI-II Neuroticism	29.00 (9.4)	26.58 (7.6)	29.10 (10.3)	0.57	
BFI-II Agreeableness	49.13 (5.7)	43.88 (6.1)	46.70 (4.4)	6.93	**
BFI-II Conscientiousness	52.07 (4.5)	46.71 (7.6)	44.90 (6.1)	7.90	**
BFI-II Openness	48.53 (5.8)	43.67 (4.5)	44.90 (8.1)	2.05	
UPPS-P Total Score	109.8 (20.5)	113.1 (21.4)	112.5 (14.2)	0.35	
UPPS-P Positive Urgency	23.1 (7.3)	23.0 (7.2)	22.5 (6.7)	0.08	
UPPS-P Negative Urgency	23.4 (7.0)	21.7 (6.3)	21.7 (3.9)	0.32	
UPPS-P Lack of Premeditation	18.3 (4.9)	20.4 (4.9)	19.6 (4.5)	2.51	
UPPS-P Lack of Perseverance	16.4 (2.8)	16.1 (3.5)	17.4 (3.4)	1.53	
UPPS-P Sensation Seeking	28.7 (6.2)	31.9 (8.2)	31.4 (9.3)	2.68	
SSS-V Total Score	12.25 (6.0)	16.04 (5.5)	15.00 (6.6)	4.41	
SSS-V Thrill Seeking	4.7 (2.2)	6.2 (2.4)	5.00 (3.5)	4.13	

SSS-V Experience Seeking	3.5 (1.6)	4.9 (1.4)	4.7 (1.6)	7.82	**
SSS-V Disinhibition	1.8 (2.1)	2.8 (2.9)	3.00 (1.4)	3.17	
SSS-V Boredom Susceptibility	2.3 (2.0)	2.0 (1.6)	2.07 (2.0)	.09	
Self-Ordered Pointing Task (SOP)					
SOP Total Errors Six Item Trials	2.4 (3.6)	1.1 (0.8)	1.1 (0.9)	1.22	
SOP Total Errors Eight Item Trials	2.9 (5.0)	1.8 (1.4)	2.5 (1.2)	2.39	
SOP Total Errors Ten Item Trials	4.1 (6.4)	3.1 (1.9)	4.4 (2.2)	0.42	
SOP Total Errors Twelve Item trials	5.5 (7.8)	4.6 (2.7)	4.8 (2.1)	0.94	
SOP Total Errors (All Trials)	14.9 (22.2)	10.6 (4.6)	11.7 (4.9)	0.85	
Trail-Making Test (TMT)					
TMT Part A # Errors	0.75 (0.78)	1.55 (1.71)	0.91 (1.22)	3.31	
TMT Part A Completion Time	52657.19 (11017.1)	58496.45 (19788.6)	52830.82 (16415.4)	1.87	
TMT Part B # Errors	2.38 (2.3)	2.00 (1.9)	1.91 (2.3)	0.49	
TMT Part B Completion Time	80812.81 (27066.8)	73383.09 (30498.1)	77767.73 (30498.1)	1.24	
Multi-Source Interference Task (MSIT)					
MSIT Congruent Trial Accuracy (proportion correct)	0.87 (.23)	0.88 (.19)	0.87 (.18)	0.79	
MSIT Incongruent Trial Accuracy (proportion correct)	0.87 (.20)	0.84 (.21)	0.80 (.22)	4.54	*
MSIT Congruent Trial Reaction Time (ms)	1021.68 (135.99)	965.23 (178.04)	963.09 (184.08)	2.71	
MSIT Incongruent Trial Reaction Time (ms)	1054.03 (143.4)	1013.10 (173.1)	1028.75 (211.3)	1.24	
MSIT Conflict Effect (Accuracy)	.00 (.07)	.03 (.06)	.07 (.08)	2.88	
MSIT Conflict Effect (RT)	32.4 (35.0)	47.9 (72.2)	65.7 (53.7)	2.84	
Attentional Network Task (ANT)					
ANT Congruent Trial Accuracy % Correct	93.4 (22.7)	94.6 (20.0)	97.2 (3.2)	2.84	

ANT Incongruent Trial Accuracy % Correct	62.5 (42.7)	76.4 (34.3)	83.8 (26.8)	1.61	
ANT Congruent Trial Reaction Time (ms)	566.24 (107.0)	581.41 (94.2)	554.21 (72.9)	0.41	
ANT Incongruent Trial Reaction Time (ms0	741.9 (232.1)	764.3 (253.7)	662.9 (129.9)	0.88	
ANT Orienting Effect	1.46 (53.4)	16.1 (33.2)	13.0 (40.2)	1.16	
ANT Conflict Effect	178.1 (147.9)	182.9 (215.9)	108.6 (77.2)	1.10	

Values represent means and, in parentheses, standard deviations. Kruskal-Wallis H-statistics: * = p<.10; ** = p<.05; ***= p<.01.

Table A7. Associations between Driver Age and Measured Driver Behaviors

	Excludes Missing and Invalid Cases			Excludes Missing and Invalid Cases + Outliers		
Task Variable	Participant Age r=	p-value	Sig	Participant Age r=	p-value	Sig
Beck Depression Inventory Total	-.22	.038	**	-.19	.071	*
Body Mass Index (BMI)	.14	.191		.13	.205	
AUDIT Total Score	-.34	.001	***	-.31	.003	***
PSQI Global Score	-.29	.005	***	-.29	.005	***
Subjective Sleep Quality	-.28	.007	***	-.28	.007	***
Sleep Latency	-.27	.010	**	-.28	.007	***
Sleep Duration	.05	.620		.04	.696	
Habitual Sleep Efficiency	-.08	.456		-.02	.840	
Sleep Disturbances	.14	.186		.14	.186	
Use of Sleep Medication	-.27	.009	***	-.22	.041	**
Daytime Dysfunction	-.28	.006	***	-.28	.006	***
Matrix Reasoning T-Score	.31	.002	***	.33	.001	***
BFI-II Extraversion	.36	.001	***	.36	.001	***
BFI-II Neuroticism	-.09	.415		-.09	.415	
BFI-II Agreeableness	.37	.000	***	.37	.000	***
BFI-II Conscientiousness	.31	.003	***	.31	.003	***
BFI-II Openness	.25	.018	**	.23	.036	**
UPPS-P Total Score	-.33	.001	***	-.33	.001	***
UPPS-P Positive Urgency	-.27	.008	***	-.27	.008	***
UPPS-P Negative Urgency	-.19	.074	*	-.19	.074	*
UPPS-P Lack of Premeditation	-.12	.235		-.12	.235	
UPPS-P Lack of Perseverance	-.13	.200		-.13	.200	
UPPS-P Sensation Seeking	-.29	.005	***	-.29	.005	***
SSS-V Total Score	-.18	.083	*	-.18	.083	*
SSS-V Thrill Seeking	-.14	.183		-.14	.183	

SSS-V Experience Seeking	.17	.101		.17	.101	
SSS-V Disinhibition	-.24	.021	**	-.22	.035	**
SSS-V Boredom Susceptibility	-.25	.017	**	-.28	.006	***
Self-Ordered Pointing Task (SOP)						
SOP Total Errors Six Item Trials	-.13	.217		-.07	.498	
SOP Total Errors Eight Item Trials	.08	.458		.16	.146	
SOP Total Errors Ten Item Trials	.07	.488		.15	.171	
SOP Total Errors Twelve Item trials	.11	.311		.18	.093	*
SOP Total Errors (All Trials)	.09	.378		.17	.116	
Trail-Making Test (TMT)						
TMT Part A # Errors	.15	.169		.16	.135	
TMT Part A Completion Time	.47	.000	***	.47	.000	***
TMT Part B # Errors	.08	.427		.09	.395	
TMT Part B Completion Time	.23	.028	**	.24	.025	
Multi-Source Interference Task (MSIT)						
MSIT Congruent Trial Accuracy (proportion correct)	.09	.422		.07	.561	
MSIT Incongruent Trial Accuracy (proportion correct)	.24	.022	**	.27	.018	**
MSIT Congruent Trial Reaction Time (ms)	.37	.000	***	.40	.000	***
MSIT Incongruent Trial Reaction Time (ms)	.31	.003	***	.34	.003	***
Attentional Network Task (ANT)						
ANT Congruent Trial Accuracy % Correct	-.03	.776		.04	.753	
ANT Incongruent Trial Accuracy % Correct	-.06	.561		.01	.966	

ANT Congruent Trial Reaction Time (ms)	.31	.003	***	.33	.003	***
ANT Incongruent Trial Reaction Time (ms0	.35	.001	***	.36	.002	***
ANT Orienting Effect	-.06	.598		-.07	.547	
ANT Conflict Effect	.28	.009	***	.22	.062	*

r-values represent Spearman Rank Order Correlations; *=p<.10; **= p<.05; ***= p<.01; if there were no outliers, values are repeated for parts 1 and 2 of the table.