



# The utility of cognitive testing to predict real world commercial driving risk

Daniel Scott<sup>a,\*</sup>, Alice Elizabeth Atkin<sup>b,c,\*</sup>, Aaron Granley<sup>a</sup>, Anthony Singhal<sup>b,c,\*</sup>

<sup>a</sup> Impirica, Canada

<sup>b</sup> Department of Psychology, University of Alberta, Canada

<sup>c</sup> Neuroscience & Mental Health Institute, University of Alberta, Canada

## ARTICLE INFO

### Keywords:

Impairment  
Commercial driving  
Risk  
Assessment  
Cognition  
Perception  
Action

## ABSTRACT

**Background:** Driving is a complex task which requires numerous cognitive and sensorimotor skills to be performed safely. On-road driver evaluation can identify unsafe drivers but can also be expensive, risky, and time-consuming. Poor performance on off-road measures of cognition and sensorimotor control has been shown to predict on-road performance in privately-licensed light vehicle drivers, but commercial drivers have not yet been studied despite such vehicles generally being larger and heavier, thus increasing risks from unsafe driving.

**Method:** Commercially-licensed truck, bus, and light vehicle drivers undertook the tablet-based Vitals cognitive screening tool, which measures reaction time, judgement, memory, and sensorimotor control, and also undertook an on-road driving evaluation using their vehicle. Accuracy and reliability of the Vitals tasks on predicting road test outcomes were investigated using a trichotomous classifier (pass, fail, borderline), and task performance was analyzed depending on vehicle type and road test outcome.

**Results:** Performance on the Vitals tasks predicted on-road performance across all vehicle types. Participants who failed their on-road evaluation also demonstrated lower success on the Judgement task, fewer correctly replicated shapes on the Memory task, and less time on-target in the Control task compared to those who passed.

**Conclusion:** Performance on cognitive and sensorimotor tasks is a good predictor of future driving performance and driver safety for commercially-licensed drivers. Regardless of vehicle type, stakeholders can use cognitive measures from the Vitals assessment to identify an increased driving risk. Use of the Vitals as a screening tool prior to on-road evaluation can benefit both drivers and evaluators.

## 1. Introduction to cognition, driving, and impairment

Driving is an important and complex task which requires various human cognitive and sensorimotor functions in order to maintain safety. For example, safe vehicle operation in traffic requires attention, memory, visual discrimination, object recognition, language, and executive functions such as problem-solving and decision-making (Canadian Council of Motor Transport Administrators [CCMTA], 2022). Impairment to any of these functions may result in reduced performance or an elevated risk of error, resulting in an increased risk of collisions and crashes. Cognitive impairment can have many levels, from mild to severe, and can be either episodic or persistent, depending on the underlying cause. Causes of impairment can include physical injury or disease, psychiatric illness, use of pharmaceutical or recreational drugs such as alcohol or cannabis, as well as situational factors such as active or passive fatigue (Saxby et al., 2013, Hancock and Desmond, 2000). Depending on the nature and severity of impairment, driving

performance can sometimes be maintained via compensatory devices or strategies, but for some drivers, the only safe solution to impairment is to cease driving, either temporarily or permanently (CCMTA, 2022).

### 1.1. Driving

Road safety is a significant public health issue all around the world. According to the World Health Organization, as of 2018, traffic fatalities were the leading cause of death in young people aged 5–29 worldwide and the eighth highest cause of death amongst all ages. However, these fatalities are not evenly distributed amongst nations, and traffic fatality rates are up to three times higher in low-income countries compared to high-income countries. In fact, unlike low-income countries, high-income countries generally have experienced a substantial decline in traffic fatality rates; since 2010, every G7 country and the European Union has seen a decline in fatality rates, with the exception of the United States (Governors Highway Safety Administration [GHSA],

\* Corresponding authors.

E-mail addresses: [dscott@impirica.tech](mailto:dscott@impirica.tech) (D. Scott), [aliceatkin@ualberta.ca](mailto:aliceatkin@ualberta.ca) (A.E. Atkin), [asinghal@ualberta.ca](mailto:asinghal@ualberta.ca) (A. Singhal).

<https://doi.org/10.1016/j.trip.2023.100783>

Received 19 October 2022; Received in revised form 10 February 2023; Accepted 12 February 2023

Available online 28 February 2023

2590-1982/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

2022; Organization for Economic Co-operation and Development [OECD], 2022). This general decline comes despite progressive aging of the population within high-income countries and a large increase in the number of seniors over the age of 75, who generally match older teenagers (18–20) for the highest rate of traffic fatalities by age group (International Transport Forum [ITF], 2020, 2021). For instance, the median age of Canadians has increased from 26.2 in 1971 to 41.1 in 2021, and the number of Canadian seniors over the age of 75 has also increased by almost 5-fold (Statistics Canada, 2021). These national differences in fatality rates and trends suggest that by identifying and addressing relevant factors, the risks of driving can be substantially mitigated, even as the number of at-risk road users increases. Many countries and cities around the world have committed to achieve precisely this via “Vision Zero” pledges which seek to eliminate traffic fatalities and injuries through changes to traffic design and public policy (Johansson, 2009; Kim et al., 2017).

All drivers share the same road infrastructure, but the vehicles that are driven can vary significantly in their design and function. Consequently, driver's licenses are typically divided into different classes, each requiring separate knowledge and on-road driving tests. There are licenses for operating a personal vehicle, such as an automobile or a motorcycle, and there are also licenses for operating commercial vehicles, such as trucks or buses. Commercial vehicles are typically larger and heavier than personal vehicles, which increases their hazard to other road users (Matsui et al., 2016; Zou et al., 2017). In general, a larger, heavier vehicle will have longer braking distances, larger blindspots, be more difficult to navigate on complex urban streets, and have greater impact energy in the event of a collision. Heavier vehicles tend to be safer for their drivers in the event of a crash, but also increase the risk of injury and death for other road users, especially pedestrians and cyclists (Anderson and Auffhammer, 2014; Li, 2012; Savage, 2013; Schmitt, 2020). Although the large majority of traffic fatalities involve personal vehicles, commercial heavy trucks have a relatively higher fatality rate per vehicle mile, and also incur extremely high costs when crashes occur (Federal Motor Carrier Safety Administration [FMCSA], 2007; Savage, 2013). In fact, in the United States, commercial truck drivers have some of the highest workplace fatality rates of any occupation (Bureau of Labor Statistics [BLS], 2021), and the nature of the occupation means that commercial truck drivers have notably high rates of health conditions that are associated with an elevated crash risk, such as sleep apnea, hypertension, and diabetes (Abu Dabrh et al., 2014). Bus crashes are very infrequent and make up only a small percentage of large vehicle crashes in the United States, but their large size and high passenger occupancy make the consequences of a crash especially serious (FMCSA, 2009). Consequently, commercial trucks and buses require a high level of driver skill and vigilance in order to be operated safely, thus necessitating special licensing above and beyond what is required to operate personal vehicles. Furthermore, personal automobiles require skill and vigilance to be operated, making driving performance a good proxy for an individual's general cognitive health and performance (CCMTA, 2022). A good driver is likely to be in good cognitive health, and a person in good cognitive health is likely to be a good driver.

## 1.2. Driving and impairment

As a complex task involving numerous cognitive and sensorimotor functions, driving is sensitive to any impairment in these functions, with the level of risk depending upon the level of impairment and its persistence. Every driver is at risk of certain forms of transient impairment, such as fatigue, but for most forms of impairment, the risk tends to be concentrated amongst certain drivers. For instance, government research in Canada and the Netherlands suggests that repeat offenders are responsible for the majority of impaired driving violations and serious crashes involving alcohol (Clermont, 2018; Goldenbeld et al., 2016). Other causes of cognitive impairment, such as neurological, cardiovascular, or metabolic diseases, reliably increase in prevalence

with increased age (Dumurgier and Tzourio, 2020; Matthews et al., 2016). For example, in the United States, approximately 0.4 % of people under the age of 75 are diagnosed annually with Alzheimer's disease, but for people over the age of 85, that number is 7.6 % (Alzheimer's Association, 2022). A diagnosis of mild cognitive impairment (MCI), which can be acquired via acute neurological injury but may also be a precursor to dementia, likewise becomes more prevalent with age, as do Type II diabetes (Sharma et al., 2016) and heart disease (Dai et al., 2021). Depending upon the disease and its severity, the resulting impairment can sometimes be minimized via pharmaceuticals and other health interventions, or it can be compensated for via changes in driving behaviour. Some neurological illnesses, such as Alzheimer's disease, however, are progressive and degenerative, meaning that drivers with these conditions will experience an irreversible deterioration in their cognitive and sensorimotor skills until it is no longer safe to drive.

Traditional crash reports suggest that the vast majority of traffic fatalities either involve impairment directly as a factor, such as when drivers are operating their vehicle under the influence of alcohol or drugs, or they can involve driver errors which are substantially more likely to occur if drivers are operating under cognitive impairment (Dingus et al., 2016). Such driver errors come in different forms and are commonly classified as either recognition errors, decision errors, or performance errors (Singh, 2015). A recognition error is when a driver fails to perceive or attend to some critical aspect of the road environment, such as a stop sign, traffic light, or pedestrian. A decision error is when a driver makes a decision that deviates from safe or legal driving practices, such as speeding, tailgating, or making unsafe maneuvers across traffic lanes. A performance error is when a driver fails to properly execute an intended driving maneuver, such as reacting slowly to an obstacle or turning the vehicle imprecisely. A narrow focus on driver error as the cause of most crashes has been criticized by the Vision Zero approach to traffic safety, which instead adopts a “systems perspective” (Kim et al., 2017) to crash analysis by assessing driver behaviour in the wider context of vehicle design, road design, and transportation policy, and by advocating for vehicles and roads which are more tolerant of error (Johansson, 2009; Schmitt, 2020). Nevertheless, it remains the case that driver errors are more likely to occur when a driver is cognitively impaired, particularly by alcohol (Compton and Berning, 2015; Dingus et al., 2016) and cannabis (Bondallaz et al., 2016; Compton, 2017; Hartman and Huestis, 2013), and that some policies to reduce impaired driving have proven successful and are globally recommended, such as lowering legal blood alcohol concentration (BAC) limits, as well as enforcement and points systems which can lead to license suspensions or criminal penalties for drivers who exceed those limits (Castillo-Manzano and Castro-Nuño, 2012; Fell and Scherer, 2017; World Health Organization [WHO], 2021).

## 1.3. License renewal and driver screening

Concerns about impaired driving mean that in addition to administering tests when drivers first acquire their license, many jurisdictions now require certain drivers to have their on-road driving skills regularly re-tested as a condition of license renewal. For instance, in the province of Alberta, Canada, drivers must pass a road test at age 75, age 80, and then every two years thereafter (Government of Alberta, 2022a). This age-based testing schedule is even stricter for commercial drivers, with such tests taking place every-five years until age 45, every-two years until age 65, and then every year thereafter. Road tests are also often required for drivers who have their licenses suspended after a drug or alcohol-related traffic offense (Government of Alberta, 2022b), and medical professionals can be empowered to recommend that a patient undertake a road test whenever impairment is suspected due to poor health (Government of Alberta, 2022c). While license renewal programs for older adults are intended to identify cognitively impaired drivers and remove them from the road, the evidence of their effectiveness is often found to be minimal (Ichikawa et al., 2015; Langford et al., 2004;

Mitchell, 2008; Rock, 1998; Vanlaar et al., 2016). Since it is undoubtedly the case that drivers who are at a high risk of crashing due to impairment or frailty should not be driving, the questionable effectiveness of currently extant license renewal programs and the broad burden that they impose on drivers and jurisdictions suggests a potential role for off-road assessments of cognitive and sensorimotor skills which are important for driving, and which can reliably recommend drivers for an additional on-road test based on individual rather than population-level risk.

One established method of driver screening is the DriveABLE Cognitive Assessment Tool (DCAT), created by Impirica. The DCAT, first introduced in 1997, includes a variety of tasks that measure cognitive ability and can be administered via computer in <30 min (Dobbs, 1997). Its revised version, Vitals, can be administered on a tablet in the same amount of time. These tasks include measures of reaction time, decision-making, working memory, and sensorimotor coordination and are conceptually related to the forms of driver error (recognition, decision, and performance errors) commonly implicated in crashes (CCMTA, 2022; Impirica, 2022). The DCAT is employed as an assessment tool in Canada by a number of healthcare systems and transportation companies, and performance on the DCAT and Vitals has shown good predictive value in several studies for whether drivers will subsequently pass or fail a road test (Dobbs, 1997, 2013; Dobbs et al., 1998; Korner-Bitensky and Sofer, 2009), particularly within groups of patients who have been diagnosed with a specific neurological condition such as dementia (Bakhtiari et al., 2020) or stroke (Choi et al., 2015).

Two recent studies from our research group will serve to illustrate the potential utility of Vitals delivered on a handheld tablet as a screening test for at-risk drivers. This test battery is comprised of tasks to test for simple reaction time, attention shifting, decision-making, memory, and sensorimotor control. In one study, we demonstrated that drivers who were identified by their physician as being at-risk for unsafe driving performed worse on the test battery relative to healthy controls (Bakhtiari et al., 2020). Furthermore, poorer performance on the test battery in healthy older adults strongly predicted subsequent failure on an on-road driving test. In a second study, we have found that Vitals performance differed in adults under the influence of cocaine, cannabis, or both cannabis and cocaine (Tomczak et al., 2022). The results showed differential sensitivity to various cognitive factors depending on which drug group they were in, and in comparison to adults not under the influence of drugs or alcohol. We made the key assumption that none of the drug influenced individuals were fit to drive. Together, these two studies indicate that drivers who are likely to pose a considerable risk to themselves or others, and who are likely to fail an on-road driving test, can be reliably identified using an off-road, tablet-based screening tool that can be administered relatively quickly, efficiently, and at low cost. We believe a reliable and widely used computer or tablet-based screening tool could significantly reduce the regulatory burden on drivers and improve the effectiveness of license renewal programs by dissociating cognitively healthy and at-risk individuals prior to a road test, with road tests reserved only for drivers at high risk of unsafe driving and test failure.

#### 1.4. Current study

In this study, we had two main purposes. First, we sought to examine performance differences on the Vitals tablet-based cognitive test battery for commercial vehicle licensing by administering it to commercially licensed truck, bus, and automobile drivers. As described in the foregoing review, these commercial vehicle types differ in many ways, and the driver demands of the different vehicles are unique. It remains an open question whether trained drivers in each of these categories have a different relationship between cognitive performance and on-road driving success. The second purpose of this study was to validate the overall predictive value of the Vitals tablet-based cognitive test battery for on-road performance across all commercial driving categories. The

strength of this approach is the large number of drivers that can be used for multivariate predictability. Truck and bus drivers pose a particular safety risk when cognitive impairment is suspected, given the large size of their vehicles and the higher level of skill required to operate them. The stricter testing schedule that commercial drivers face when seeking to retain their license also means that the benefits of a computer-based screening test may be especially high for these drivers, relative to drivers who are only licensed to operate personal vehicles. Thus, in this study, we had commercial truck, bus, and small vehicle drivers perform the Vitals cognitive test battery and also perform a vehicle-specific on-road test. We examined their performance on each dependent measure of the cognitive tests and used the aggregate data to explore the predictability of cognitive task performance to overall commercial driving success.

## 2. Methods

### 2.1. Participants

We compared data from three separately recruited populations associated with different commercial vehicle types. Table 2 contains the available demographic data for each of the groups. The first group was Heavy vehicles, consisting of 415 adult participants (mean age = 41.7 yrs., SD = 13.0). The second group was the Light vehicles, including car, van, and light commercial truck drivers consisting of 1592 adult participants (mean age = 36.95 yrs., SD = 10.7). The third group was Trucking, consisting of 160 adult participants (mean age = 43.69 yrs., SD = 12.57). Due to procedural constraints, we were unable to accurately record the gender of the participants. The median time spent on the on-road driving assessment was 64 min, although the time distributions varied across organizations. Differences in time were best explained by the unique road constraints required for different vehicle types. For example, large trucks were limited to driving on designated trucking roads, but small vehicles were not.

This study was conducted according to the Declaration of Helsinki (1996). It was approved by the University of Alberta Health Research Ethics Board and performed in compliance with relevant laws and institutional guidelines. All participants gave informed consent to have their data collected for this study. Completion of the Vitals assessment and on-road assessment was linked to each participant's job duties or was performed as a condition of their hiring by an organization, but participants had the option of completing the evaluations whilst refusing consent to have their data included in this study. The study took approximately 90 min to complete – 30 min for the Vitals and 60 min for the on-road assessment.

### 2.2. Procedure

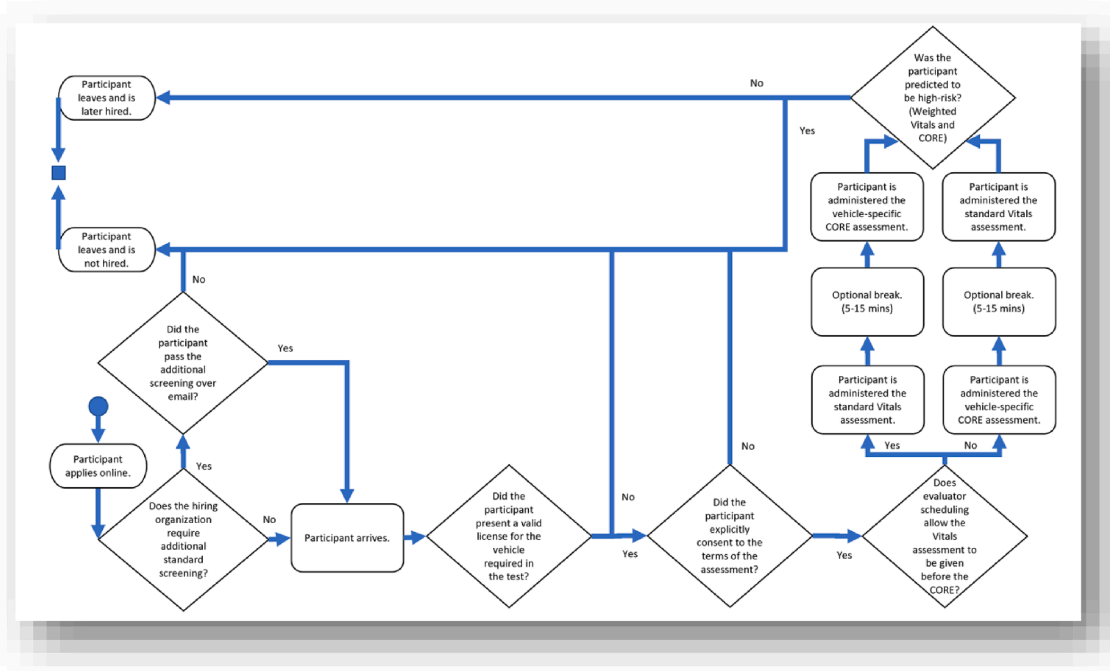
The Vitals cognitive screen was based on decades of cognitive science research and designed to test the primary cognitive, perceptual, and action-related measures intimately related to driving performance (Bakhtiari et al., 2020). Its tasks include reaction time (RT), decision-making (Judgement), Memory, and sensorimotor control (Control). The Commercial On-Road Evaluation (CORE) was designed to test driving ability judged by a trained instructor for each of the vehicle types. Fig. 1 shows the process employed for data collection for both the Vitals and CORE assessments.

### 2.3. Tasks

#### 2.3.1. Cognitive tasks (vitals)

The four manually controlled, tablet-based tasks included in this study were designed and developed by Impirica to run as a continuous ordered series. These four tasks included:

- 1) Reaction Time (RT): participants press a button as fast as possible after a visual cue. The cue can be spatially congruent with the button



**Fig. 1.** Unified Modelling Language (UML) diagram with detail of the participant screening process and decision tree for the CORE/Vitals pre-hiring procedure. Participants who were being assessed as part of a performance evaluation, for periodic wellness, or after an at-work incident, may have been subject to additional screening related to the continuation of their duties.

location or incongruent. Thus, this task also tests for attention switching.

- 2) Judgement, i.e., decision-making: participants press a 'Go' button after a visual cue while avoiding moving obstacles. A 'Stop' button is given as an additional control to avoid obstacles. There are two stages. Stage one has one set of obstacles and the second stage has two sets of obstacles. There are ten trials in the first stage and 20 trials in the second stage.
- 3) Memory: participants draw a previously presented geometric shape with their finger. There are four stages, with four trials in each stage. The number of shapes to recall, shape complexity, and the duration of an intervening mask increases as the stages progress.
- 4) Control, i.e., bi-manual sensorimotor control: participants follow a target circle using the iPad as if it was a steering wheel to avoid fixed and unpredictable moving obstacles. There are four stages, where speed increases as the stages progress.

Vitals uses reliable measures of cognitive processes known to be associated with safe driving and are predictive of on-road performance in cognitively impaired drivers (Dobbs, 1997, 2013; Dobbs et al., 1998; Bakhtiari et al., 2020). The RT task is a modified version of an attentional shifting task (Monsell, 2003; Robbins, 2007), while the Judgement task is related to other tasks of spatial judgement. The Memory task is a modified version of the Corsi block test (Kessels et al., 2000). The Control task is chosen because it requires similar sensorimotor coordination to that of on-road driving.

In designing Vitals, tests were selected to cover most of the required cognitive domains needed for driving (CCMTA, 2022; Dobbs, 1997); divided attention (the ability to attend to two or more stimuli at the same time, evaluated in bi-manual sensorimotor task), selected attention (the ability to selectively attend to one or more important stimuli while ignoring competing stimuli, evaluated in reaction speed and decision-making tasks), sustained attention (the capacity to maintain attentional activity over a period of time, evaluated by bi-manual sensorimotor task), short-term memory (the temporary storage of information that is currently being processed in a person's mind, evaluated by

memory task), working memory (the ability to manipulate information with time constraints/taking in and updating information to solve problems, evaluated by all four tasks), complex reaction time (the time taken to respond differentially to two or more stimuli or events, evaluated by reaction speed and decision-making tasks), tracking (the ability to visually follow a stimulus that is moving or sequentially appearing in different locations, evaluated by bi-manual sensorimotor task), visuo-spatial abilities (processes dependent on vision such as the recognition of objects, the ability to mentally rotate objects and determination of relationships between stimuli based on size or color, evaluated by decision-making task), executive functioning (capabilities that enable an individual to successfully engage in independent, purposeful, and self-serving behaviors, evaluated by all four tasks), and visual information processing (the processing of visual information beyond the perceptual level, evaluated by all four tasks).

The Vitals cognitive test battery is administered by an evaluator that is trained and certified. Each task includes a demonstration and practice. If the administrator evaluates that an individual does not comprehend a task during its demonstration, the demonstration is repeated to aid comprehension but not to improve performance. If the administrator evaluates that the participant comprehends the task, but performs poorly, the demonstration is not repeated. After the demonstration is complete, the administrator can answer the participant's questions with the provided standard responses from a manual.

### 2.3.2. Commercial On-Road Evaluation (CORE)

The CORE is a standardized on-road driving assessment with a trained driving instructor. The evaluation takes approximately 60 min to complete and is based on a scientifically developed on-road evaluation designed to test for a decline in the cognitive skills necessary for safe driving (Berndt et al., 2007; Dobbs et al., 1998; Kowalski and Tuokko, 2007). The CORE is a modification of the DriveABLE On-Road Evaluation (DORE), which was previously reported in Bakhtiari et al. (2020). The CORE assesses the same primary error categories as the DORE but includes additional categories depending on each participant's vehicle type. This is due to physical differences in vehicle size and controls,



including brakes, maneuverability, blind spots, and necessary visual scanning. The total difference in errors between vehicle types is 4 % (three differences across 70 error categories).

## 2.4. Vitals dependent measures

For each participant, dependent measures were extracted from the Vitals cognitive test battery. Table 1 contains a list of dependent measures from each task. They are as follows. From the RT task, the average reaction times (Reaction time) were extracted. From the Judgement task, the percentage of trials that participants started early (% Premature start), the percentage of trials successfully passing the moving obstacles (% Success), the number of times pressing the ‘Start’ button (Start count), the average time taken to press “Go” after the visual cue (Reaction time), and the average number of obstacles that passed after the visual cue and before finishing passing the obstacles (Obstacle count). From the Memory task, duration of task completion (Duration) and the percentage of correct shape retrieval (% Correct shape) were extracted. From the Control task, the percentage of time inside the target (% Time inside target) and the percentage of times that surprise obstacles were avoided (% Surprise obstacles avoided) were extracted. All dependent measures have their non-linear effects of age removed prior to analysis and integration into models. Age effects were explained and removed using a polynomial function fit to 70 % of the training dataset (940 of 1343 training samples) and tested for stability on the remaining 30 % (403 of 1343 training samples) used for model testing within the training dataset. Similarly, the analysis performed on the validation dataset removes age effects but only uses the function estimate’s fit on the training dataset (Table 9).

## 2.5. Statistical analyses

To validate the Vitals assessment’s power to predict the CORE pass/fail outcome, we evaluated the performance of our logistic regression model across the three distinctly identifiable vehicle categories (Light, Heavy, and Trucking). All the model’s parameters were fit on a training dataset prior to the validation dataset collection. A logistic model’s decision cutpoint can be calibrated to balance the diagnostic tradeoffs of misclassifying positive and negative outcomes. When there are asymmetric consequences between correct and incorrect predictions, two cutpoints (trichotomous) can be used instead of one (dichotomous) to limit decision sensitivity to within a certain threshold of confidence. If a predicted probability of driving risk is between an upper and lower trichotomous threshold, it is set aside as an undefined decision. This protects all stakeholders from downside when only under-confident predictions are available from the model. The Vitals assessment was

**Table 2**

On-road pass/borderline/fail sample sizes for drivers of heavy vehicles, light vehicles, and trucks in the validation dataset (n = 2167).

	Driving Pass	Driving Borderline	Driving Fail	N	Age (M ± SD)
Heavy	273	65	77	415	41.7 ± 13.0
Light	919	346	327	1592	36.95 ± 10.7
Trucking	99	22	39	160	43.69 ± 12.57

designed to address real-world driving risks and must minimize the potential hazard shared by the individual and the traffic system (Bakhtiari et al., 2020). To address this, we applied a set of trichotomous thresholds which are fit to maximize both positive predictive value (likelihood of on-road failure if predicted to be high-risk on Vitals), and negative predictive value (likelihood of on-road pass if predicted to be low-risk on Vitals) measures prior to the collection of the validation dataset (Table 8). The cutpoints were optimized on the training dataset (n = 1343) using the pseudo minimization objective ( $FPR \times \text{“false positive penalty”} + FNR \times \text{“false negative penalty”}$ ). To define our cutpoint objective, distinct cases of misclassification were first identified by industry experts (e.g. participant received very high Vitals score but very good on-road performance) in the training dataset and assigned a penalty. The penalties quantified negative consequences across dimensions of participant mental hazard (false positive) and driving accident cost (false negative). Since the objective’s penalties are human labeled, discrete, and non-differentiable, a direct search optimization method is advised (Silva et al., 2018). The Nelder-Mead simplex method was selected because of its usefulness for non-differentiable unconstrained objective optimization (Silva et al., 2018). Alternatives were evaluated as the sum of distinct misclassifications multiplied by the rate of their respective occurrence across the training dataset for a given set of cutpoints. Alternatives that resulted in undefined decision rates larger than 20 % were discarded. The final trichotomous cutpoint solution which minimized the objective was 0.48 and 0.68 for the low and high cutpoints, respectively. This set of trichotomous thresholds was used for the evaluation of the validation dataset (n = 2167) and was not refit once the validation data collection began. Demographic data of the sample used to generate the classifier model are shown in Table 10. A variety of data was used, including commercial, medical, judicial, and academic sources (Table 11).

To evaluate the influence of observation heterogeneity on the explainability of a driving outcome model, a mixed effect logistic model was implemented on the validation data (n = 2167), (Bates et al., 2015). This model incorporated random intercepts for each of the six Assessment Purpose groups to account for the variation in the data among these groups. For comparison, a standard logistic model was also modeled, without random effects. Although the fixed effect coefficients of the logistic model parameters cannot be shared due to the data privacy agreement from Impirica, the p-values for each fixed effect in both models were calculated and are presented in the results section. By comparing the p-values, we can determine if the random effects within each group significantly explain more about the problem than the variables themselves. It would have been preferable to model the random effects for each participant, but this was not possible because each participant only had one sample. A single sample for each random intercept would result in a saturation of random effects within the mixed model. The implementation of a mixed effect logistic model and comparison with a standard logistic model provides valuable insight into the influence of observation heterogeneity on the explainability of the model for predicting outcomes on a standard driving test, despite the limitations in presenting the full details due to trade secret restrictions.

**Table 1**

Tasks and dependent measures from Vitals.

Task	Measure	Description
Reaction Time	Reaction time	Average reaction time to press “Stop” button after the visual cue
Judgement	% Premature starts	Percentage of releasing “Start” button before the visual cue
	% Success	Percentage of successfully navigating through the moving obstacles
	Start count	Average number of times “Start” button pressed
	Reaction time	Average time taken to press “Go” after the visual cue
Memory	Obstacle count	Average number of passed obstacles, prior to engagement
	Duration	Average time to complete task
	% Correct shape	Percentage of shapes drawn correctly
Control	% Time inside target	Percentage of time inside the target
	% Surprise obstacles avoided	Percentage of surprise obstacles avoided

### 3. Results

#### 3.1. Descriptive Statistics

Table 2 presents on-road test outcomes for each driving group (Light, Heavy, and Trucking). Based on results from the standard on-road evaluation test, drivers who drove safely or dangerously were grouped as pass or fail, respectively. Participants whose driving errors were toward the upper end, but did not exceed, the range for safe drivers are categorized as borderline. Table 3 presents the mean, standard deviation, and median scores on each measure of the Vitals tasks for the three driving groups. Table 4 presents data regarding the purpose of assessment for participants within the three driving groups. The majority of participants in each driving group were recruited to the study as part of a pre-hire evaluation for prospective employees, however, the Heavy vehicle group did include a disproportionate number of participants who were assessed as part of a performance evaluation or as part of a return-to-work process following a workplace incident. Individuals that were unable to complete a task for any reason were excluded from the task's analysis.

#### 3.2. Group comparisons

##### 3.2.1. Full sample

As seen in Table 5 and Table 6, participants were divided into pass and fail groups. Performance on each measure of the Vitals tasks was compared between groups. Participants who failed their on-road evaluation were significantly less successful on Stage 2 of the Judgement task, performed fewer false starts on Stage 2 of the Judgement task, replicated fewer shapes on the Memory task, and spent less time within the target circle on the Control task.

On each measure of the Vitals tasks, participant performance was compared depending on their driving group and whether they passed or failed the on-road evaluation. For the RT task (Fig. 2), light vehicle drivers who passed the test had a significantly faster average RT than light vehicle drivers who failed. No differences were observed for truck or heavy vehicle drivers depending on if they passed or failed their test.

On Stage 1 of the Judgement task, light vehicle drivers who passed the on-road test made significantly fewer premature starts (Fig. 3) and were more successful (Fig. 4) than drivers who failed. No differences were observed for truck or heavy vehicle drivers depending on if they passed or failed their test. No differences in average RT or the number of obstacles passed prior to engagement were observed for any driving group, depending on if they passed or failed. On Stage 2 of the task, light vehicle drivers who passed the on-road test were significantly more successful (Fig. 5) than light vehicle drivers who failed and made more starts and stops while avoiding obstacles (Fig. 7). Truck drivers who passed the on-road test were also significantly more successful (Fig. 5)

**Table 4**

Purpose of assessment for participants within each driving group in the validation dataset (n = 2167).

Assessment Purpose	Heavy Vehicles		Light Vehicles		Trucking	
	n	Proportion	n	Proportion	n	Proportion
Other	5	0.012	4	0.003	2	0.013
Performance Evaluation	70	0.169	148	0.093	3	0.019
Periodic Wellness	4	0.01	7	0.004	2	0.013
Post-Incident	17	0.041	10	0.006	2	0.013
Pre-Hire	253	0.61	1418	0.891	149	0.931
Return to Work	66	0.159	5	0.003	2	0.013
Total	415		1592		160	

**Table 5**

Bonferroni-adjusted *t*-test comparison of means by on-road test outcome (fail vs pass) in the validation dataset (n = 2167).

Variables	fail (n)	pass (n)	statistic	df	p	p. adj	sig
Reaction Time Stage 2	443	1724	1.41	601.88	0.16	1	ns
Judgement Premature Starts Stage 1	443	1724	1.94	611	0.05	0.68	ns
Judgement Premature Starts Stage 2	443	1724	1.61	575.22	0.11	1	ns
Judgement Success Stage 1	443	1724	-2.44	605.78	0.01	0.2	ns
Judgement Success Stage 2	443	1724	-6.09	530.93	0.00	0	****
Judgement Reaction Time Stage 1	443	1724	0.3	668.45	0.76	1	ns
Judgement Reaction Time Stage 2	443	1724	2.25	642.37	0.02	0.32	ns
Judgement Obstacle Count Stage 1	443	1724	0.26	663.46	0.79	1	ns
Judgement Start Count Stage 2	443	1724	-3.22	698.74	0.00	0.02	*
Memory Duration	443	1724	2.83	660.61	0.00	0.06	ns
Memory Correct Shapes	443	1724	-4.3	647.47	0.00	0	***
Control Time Inside the Circle	443	1724	-6.02	713.84	0.00	0	****
Control Surprise Obstacles Avoided	443	1724	-0.55	694.12	0.58	1	ns

**Table 3**

Mean and median performance on Vitals tasks by driving group (Heavy, Light, and Trucking) in the validation dataset (n = 2167).

Vehicle Type	Heavy				Light				Trucking			
	N	Mean	SD	Median	N	Mean	SD	Median	N	Mean	SD	Median
Reaction Time Stage 2	413	0.518	0.061	0.517	1585	0.527	0.068	0.519	158	0.505	0.08	0.502
Judgement Premature Starts Stage 1	413	0.044	0.082	0.024	1590	0.046	0.08	0.024	159	0.049	0.113	0.021
Judgement Premature Starts Stage 2	413	0.06	0.06	0.025	1590	0.057	0.059	0.025	159	0.051	0.051	0.024
Judgement Success Stage 1	413	0.892	0.157	0.961	1589	0.882	0.161	0.96	158	0.872	0.199	0.961
Judgement Success Stage 2	413	0.791	0.111	0.806	1589	0.785	0.112	0.804	159	0.77	0.142	0.803
Judgement Reaction Time Stage 1	413	2.26	1.483	1.779	1589	2.466	1.508	2.005	158	2.25	1.545	1.788
Judgement Reaction Time Stage 2	413	2.677	0.949	2.395	1589	2.838	1.064	2.57	159	2.661	0.969	2.458
Judgement Obstacle Count Stage 1	413	0.959	0.623	0.865	1589	1.077	0.647	0.889	158	0.985	0.693	0.864
Judgement Start Count Stage 2	413	1.852	0.261	1.852	1589	1.887	0.299	1.852	159	1.904	0.408	1.853
Memory Duration	413	116.142	23.575	114.927	1590	119.324	24.1	116.469	159	119.22	26.53	115.882
Memory Correct Shapes	413	0.549	0.17	0.556	1590	0.535	0.164	0.533	159	0.499	0.174	0.527
Control Time Inside the Circle	413	0.422	0.108	0.424	1590	0.375	0.113	0.377	159	0.419	0.1	0.416
Control Surprise Obstacles Avoided	413	0.908	0.077	0.921	1590	0.916	0.075	0.938	159	0.879	0.099	0.901

**Table 6**

Wilcoxon comparison of medians by on-road test outcome (fail vs pass), using Holm adjustment in the validation dataset (n = 2167).

Variables	fail (n)	pass (n)	statistic	p	p. adj	sig
Reaction Time Stage 2	443	1724	399,745	0.06	0.39	ns
Judgement Premature Starts Stage 1	443	1724	391,462	0.33	1	ns
Judgement Premature Starts Stage 2	443	1724	394,198	0.23	1	ns
Judgement Success Stage 1	443	1724	363,206	0.17	1	ns
Judgement Success Stage 2	443	1724	319058.5	0	0	****
Judgement Reaction Time Stage 1	443	1724	381,691	0.82	1	ns
Judgement Reaction Time Stage 2	443	1724	405,450	0.03	0.2	ns
Judgement Obstacle Count Stage 1	443	1724	378,836	0.99	1	ns
Judgement Start Count Stage 2	443	1724	345,039	0	0.03	*
Memory Duration	443	1724	409,045	0.01	0.12	ns
Memory Correct Shapes	443	1724	333888.5	0	0	****
Control Time Inside the Circle	443	1724	312,453	0	0	****
Control Surprise Obstacles Avoided	443	1724	370,326	0.4	1	ns

and had a faster average RT (Fig. 6) than truck drivers who failed. No differences were observed for heavy vehicle drivers on any measure depending on if they passed or failed their test. No differences were observed in the number of premature starts for any driving group.

On the Memory task, light vehicle and truck drivers who passed the on-road test performed the task faster (Fig. 8) and successfully replicated more shapes (Fig. 9) than light vehicle and truck drivers who failed the test. No differences were observed for heavy vehicle drivers on either measure depending on if they passed or failed their test.

On the sensorimotor Control task, light and heavy vehicle drivers who passed the on-road test spent significantly more time within the target circle (Fig. 10) compared to light and heavy vehicle drivers who failed, respectively. No differences were observed for truck drivers depending on if they passed or failed their test. No differences in the number of surprise obstacles hit were observed for any driving group, depending on if they passed or failed.

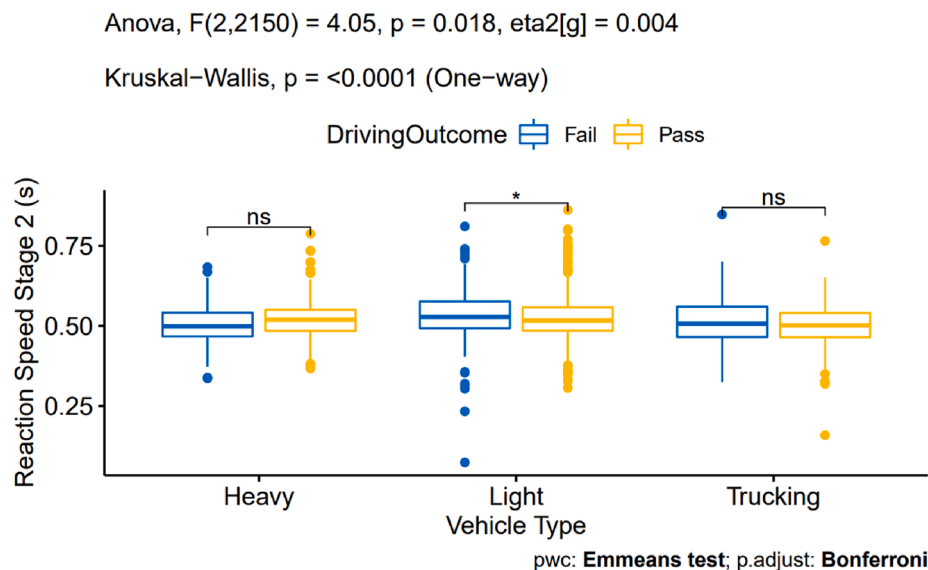
### 3.2.1. Return-to-work group excluded

Some of our participants were recruited as part of a return-to-work procedure following an at-work safety incident, and these participants were disproportionately represented in the Heavy vehicle group. In our study, there were 66 return-to-work participants in the Heavy vehicle group, compared to five and two such participants in the light and trucking groups, respectively (Fig. 2d). Return-to-work participants were highly successful at passing their on-road test despite poor performance on the Vitals tool relative to participants who were recruited for other reasons, such as pre-hire screening, performance evaluations, or routine wellness checks. There were few significant pairwise comparisons between the Heavy vehicle drivers who passed or failed their on-road test compared to the other two driving groups, and we speculated that the over-representation of return-to-work participants within the Heavy vehicle group could potentially be obscuring such statistical differences. We therefore re-ran the univariate analyses from the previous section with the return-to-work participants excluded from the analysis. Table 7 shows the new sample sizes for each driving group with the return-to-work group excluded, along with the mean and median for each measure.

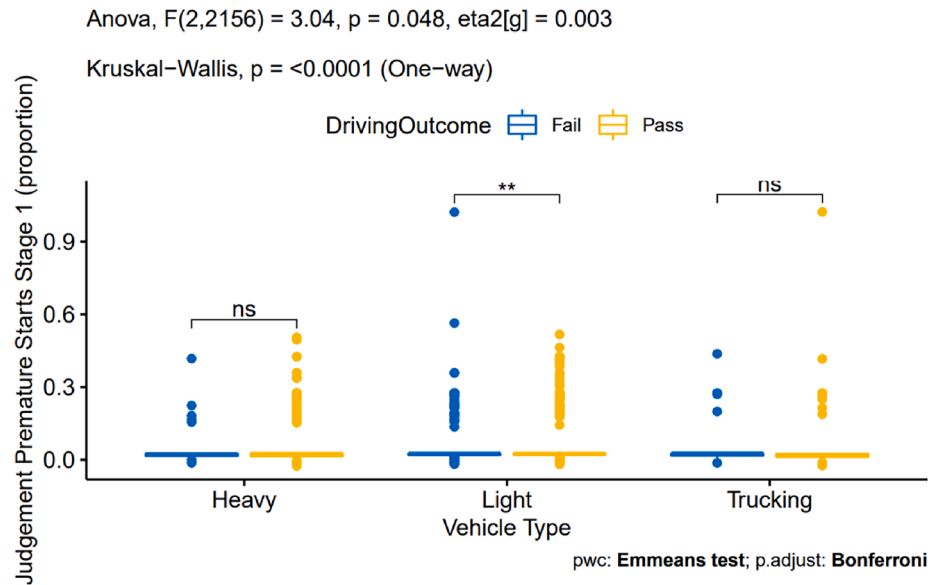
Excluding the return-to-work group from each ANOVA resulted in the outcome of several tests changing from statistically significant to non-significant, compared to the full sample. Such tests included the RT task (Fig. 3), the proportion of premature starts on the Judgement task in the second stage, trial duration on the Memory task (Fig. 9), and proportion of successfully replicated shapes on the Memory task (Fig. 10). For these four measures, excluding the return-to-work group meant that there were no longer any significant differences in Vitals performance depending upon driving group and on-road test outcome. One pairwise comparison was also affected within the Trucking group, with the proportion of successfully replicated shapes on the Memory task (Fig. 10) no longer differing depending upon on-road evaluation outcome. Most importantly, however, excluding the return-to-work participants did not affect the significance of any pairwise comparison within the Heavy vehicle group. Therefore, excluding the return-to-work participants did not confirm our speculation as to why there are few statistical differences in Vitals task performance between Heavy vehicle drivers who pass or fail their on-road test (see Table 10 and Table 11).

### 3.3. Predictability of Vitals to distinguish between safe and unsafe drivers

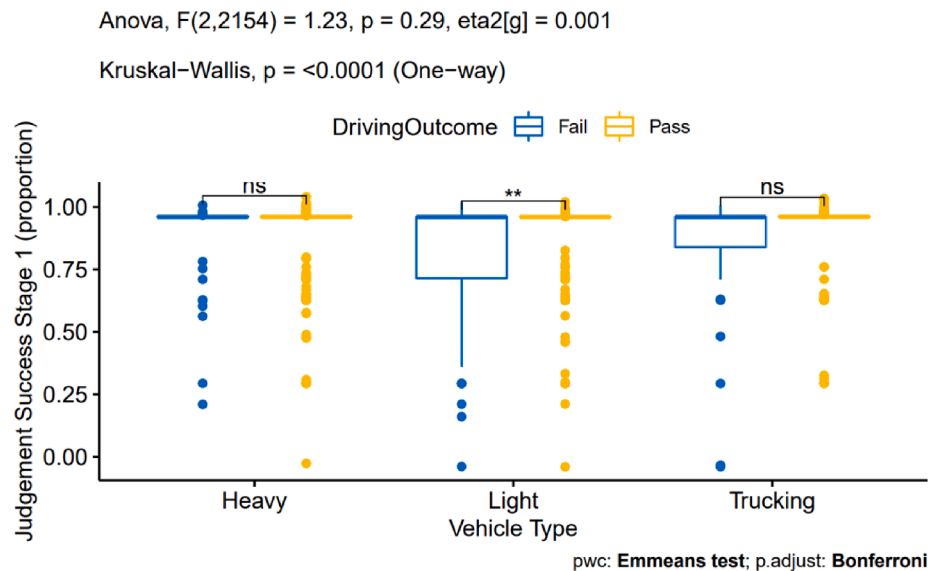
Vitals data from this study's participants was inputted into the model



**Fig. 2.** Group means and differences comparing reaction time on the RT task depending on driving group and on-road test outcome. Light vehicle drivers who failed the road test were slower than light vehicle drivers who passed.



**Fig. 3.** Group means and differences comparing proportion of total starts that were premature on Stage 1 of the Judgement task, depending on driving group and on-road test outcome. Light vehicle drivers who passed the test produced fewer false starts than light vehicle drivers who failed as a proportion of starts.



**Fig. 4.** Group means and differences comparing trial success rate on Stage 1 of the Judgement task, depending on driving group and on-road test outcome. Light vehicle drivers who passed the on-road test were more successful than light vehicle drivers who failed.

to predict whether participants would pass or fail their road test. Each participant in the study was assigned a score, calculated as the model's predicted probability of their class being 1, i.e., "impaired",  $Score = 100P(Y = 1)$ . Likelihood of participants failing their on-road evaluation depending on their Vitals score is depicted in Fig. 11. High-scoring participants were much more likely to fail the road test than low-scoring participants, showing the ability of the model to identify at-risk drivers on the basis of overall task performance.

### 3.4. Mixed effects logistic regression model

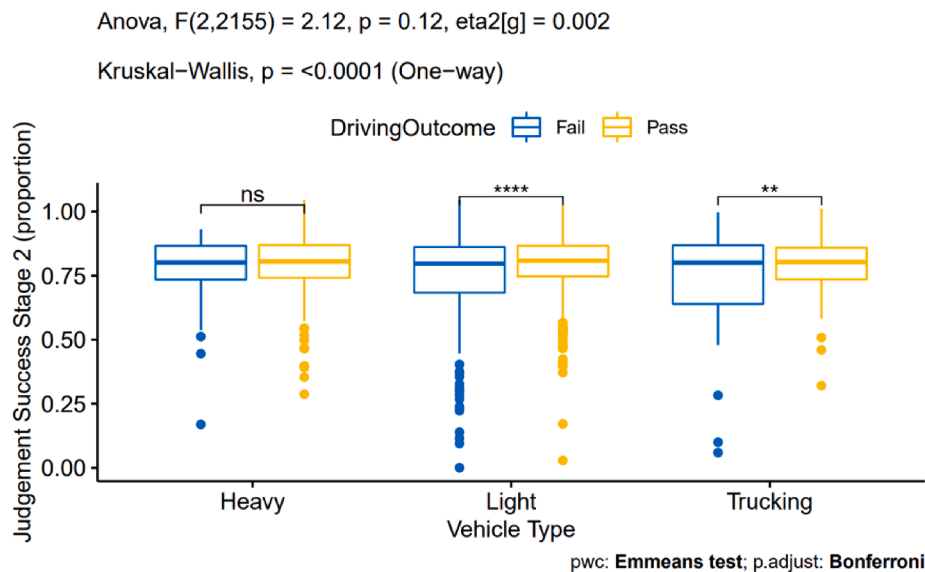
In the first model (Table 12), we used a logistic regression to examine the relationship between the response variable "Driving Outcome" and the predictor variables "Judgement Success Stage 2", "Judgement Go Count Stage 2", "Memory Correct Shape Count", and "Control Time Inside The Circle." The model was fit using the glm function in R with a

binomial distribution and a logit link function (Bates et al., 2015). The residual deviances ranged from  $-2.351$  to  $3.522$  and all predictor variables were found to be statistically significant with  $p$ -values  $<0.05$ . For the first logistic model with fixed effects only, the residual deviance was  $5738.4$  on  $2145$  degrees of freedom and the AIC was  $5225.7$ .

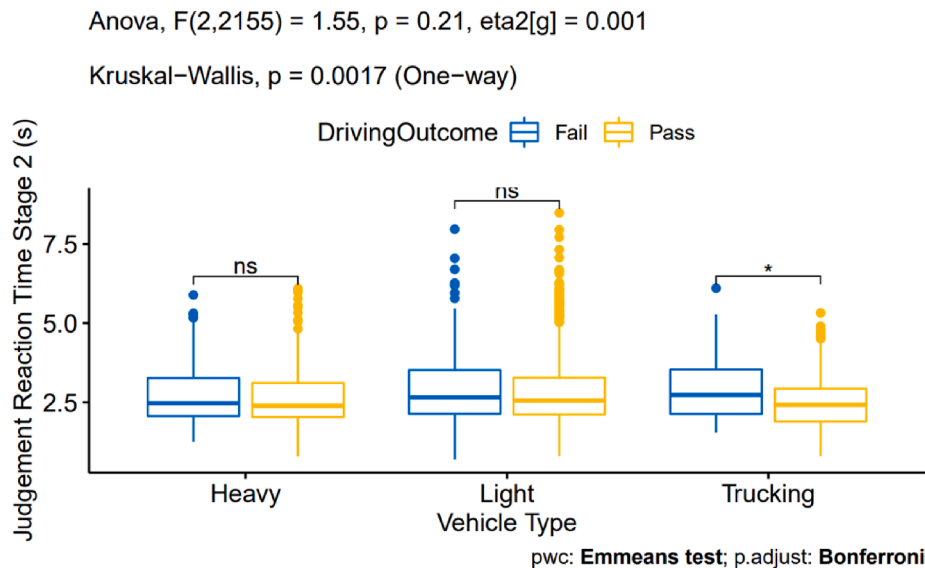
In the second model (Table 13), we added an additional random effect term for the Assessment Purpose group to the predictor variables used in the first model. This resulted in a mixed-effects logistic regression, fit using the glmer function in R with a binomial distribution and a logit link function. The AIC of the model was  $5712.9$  and the residual deviance was  $5700.9$  on  $2144$  degrees of freedom. The fixed effects estimates showed that all predictor variables were significant, with  $p$ -values  $<0.05$ .

It is worth mentioning that we selected the four predictor variables based on the results of an ANOVA test, where they were found to be significant. No regularization was used in fitting the fixed effects or





**Fig. 5.** Group means and differences comparing success rate on Stage 2 of the Judgement task, depending on driving group and on-road test outcome. Light vehicles as well as truck drivers who passed the on-road test were significantly more successful than drivers who failed.



**Fig. 6.** Group means and differences comparing reaction time on Stage 2 of the Judgement task, depending on driving group and on-road test outcome. Truck drivers who failed the on-road test were slower than truck drivers who passed.

mixed effects logistic models. The difference in AIC between the first and second models (5225.7 and 5712.9 respectively) suggests that adding the random effect term for Assessment Purpose group improved the model's fit, although the improvement may be relatively small.

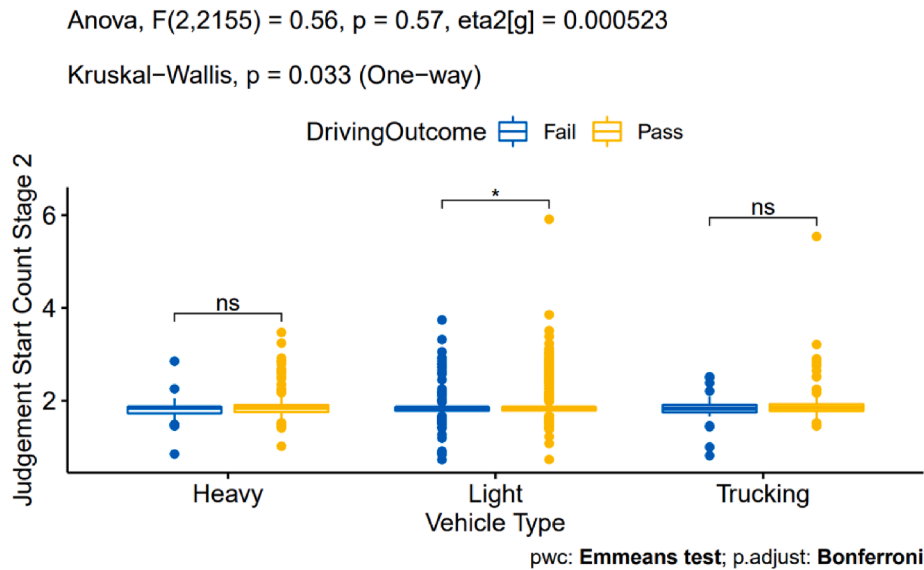
#### 4. Discussion

In this study, we sought to examine performance on the Vitals task battery in a large sample of commercial drivers and compare across three vehicle types. The Vitals tool consists of four tests of reaction time, decision-making and judgement, memory, and sensorimotor control, and has been previously found to predict on-road driving performance in a group of healthy older adults (Dobbs et al., 1998), patients with dementia (Bakhtiari et al., 2020; Dobbs, 1997) or stroke (Choi et al., 2015) and in drivers under the influence of cannabis or cocaine (Tomczak et al., 2022). In this study, we found that overall Vitals performance was associated with passing an on-road commercial driving test, as were

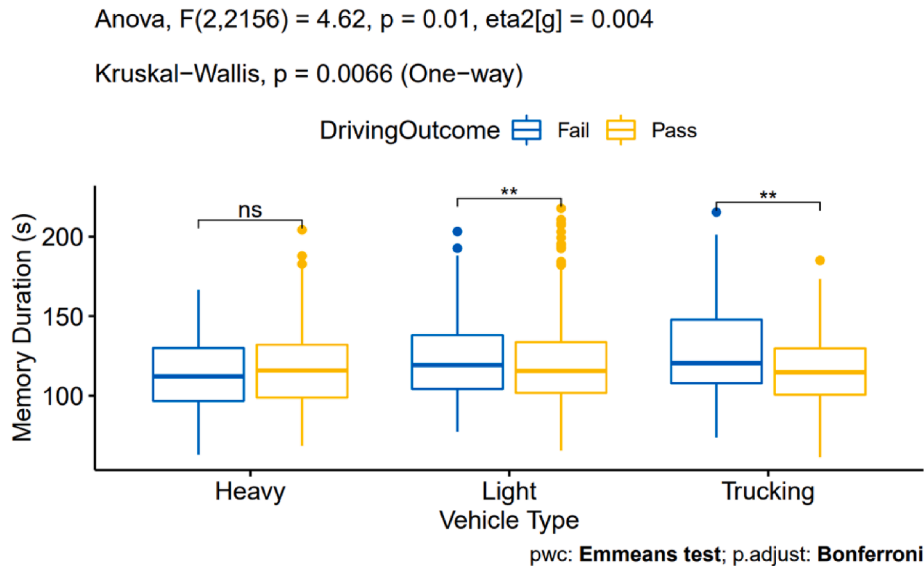
certain cognitive measures, most notably simple RT, memory performance, and time-on-target in the Control task. This set of findings suggests that there is a strong relationship between these tablet-based tasks and real-world driving. Few differences were observed between the driving groups, which may be because all three groups consisted of professionally trained drivers, and because truck and heavy vehicle drivers are highly likely to also be light vehicle drivers due to operating a personal vehicle. Greater sensitivity in the univariate analyses was associated with the light vehicle group, likely because the sample size for this group was considerably higher than the other two groups.

##### 4.1. Reaction time

The Reaction Time task on the Vitals tool requires participants to observe two or three red lights flashing in sequence, followed by a flashing green light, and respond as quickly as possible once the green light is presented. This task conceptually resembles waiting for a traffic



**Fig. 7.** Group means and differences comparing the number of stops and starts to avoid obstacles on Stage 2 of the Judgement task, depending on driving group and on-road test outcome. Light vehicle drivers who passed the on-road test had more stops and starts than drivers who failed.



**Fig. 8.** Group means and differences comparing trial duration on the Memory task, depending on driving group and on-road test outcome. Light vehicle drivers who failed the on-road test were slower to replicate the shapes than light vehicle drivers who passed. Truck drivers who passed the on-road test were faster to replicate the shapes than truck drivers who failed.

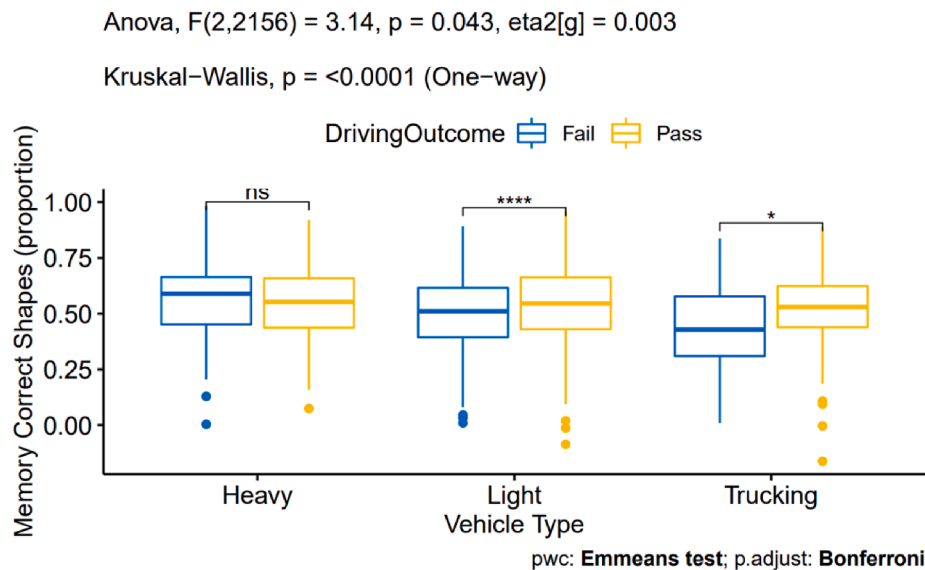
light to change before accelerating, and requires participants to engage in visual discrimination, sustained attention, response inhibition, and reaction time, among other neurocognitive abilities (Dobbs, 1997; Dobbs et al., 1998). Predictive modeling showed that faster reaction times were effective at specifying individuals who would go on to pass their on-road test, while univariate ANOVA also distinguished pass and fail drivers within the light vehicle group.

Reaction time is typically assumed to represent the speed (Pachella, 1974; Teichner, 1954) or stages of mental processing (Treisman, 1991). Slowed reaction times are often seen in individuals who have experienced a head injury (Balaban et al. 2016; Stuss et al., 1989), are physically frail (Robertson et al., 2014), or are experiencing pathological aging processes, such as mild cognitive impairment and Alzheimer's Disease (Gorus et al., 2008). Reaction time is an important aspect of driving, as drivers are often required to respond quickly to visual stimuli such as traffic lights, obstructions and debris, or other road users. A

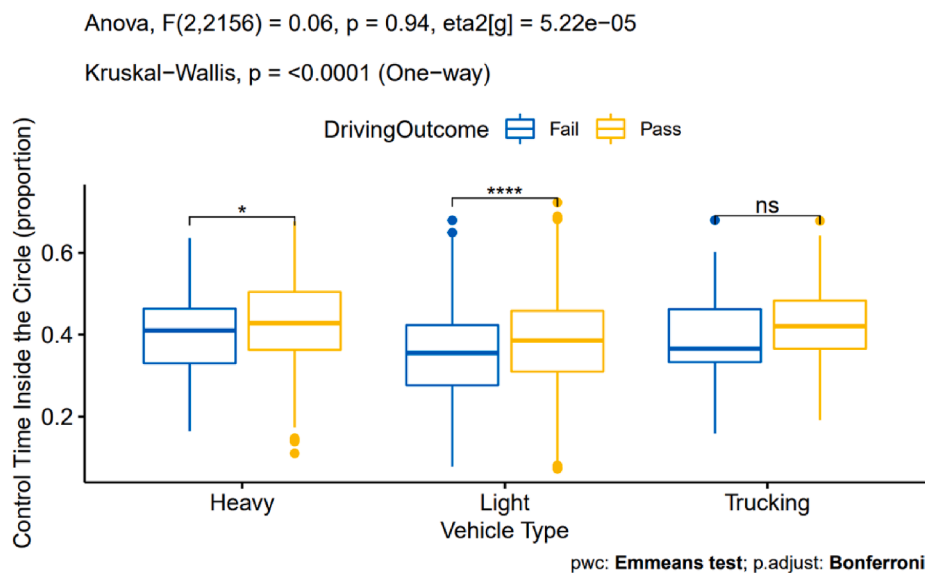
slower reaction time can translate into longer stopping distances and greater momentum at crash impact, with potentially serious or fatal consequences. Therefore, it seems unsurprising that reaction time predicts whether a driver will pass or fail a road test.

#### 4.2. Decision-making

The Judgement task requires participants to control a virtual icon and safely cross it from one side of the screen to the other using a Start/Stop button without contacting any obstacles. The task consisted of two levels of difficulty: In Stage 1, participants had to navigate across one lane of "traffic flow," which consisted of rising or falling blocks separated by navigable gaps, while in Stage 2, participants had to navigate across two lanes of flow moving in opposite direction to one another. This task conceptually resembles the act of navigating a vehicle across one or two lanes of traffic and captures several executive functions such



**Fig. 9.** Group means and differences comparing proportion of total shapes that were successfully drawn on the Memory task, depending on driving group and on-road test outcome. Light vehicle drivers who passed the on-road test drew more successful shapes than light vehicle drivers who failed. Truck drivers who passed the on-road test drew more successful shapes than truck drivers who failed.



**Fig. 10.** Group means and differences comparing proportion of total time on the sensorimotor Control task that was spent inside the target circle, depending on driving group and on-road test outcome. Light vehicles as well as bus drivers who passed the on-road test spent more time inside the circle than light vehicle and bus drivers who failed.

as attention, decision-making, and reaction time, as well as sensorimotor abilities (Dobbs, 1997; Dobbs et al., 1998).

Contrary to the RT task, reaction time in the Judgement task was not associated with on-road test outcome. The Judgement task is more complex than the RT task, as participants respond to a sequence of stimuli rather than a single stimulus. Depending on the location of stimuli at trial onset, a faster reaction time may in fact be contrary to successful task performance and lead to a collision, whereas waiting for a gap in the flow of stimuli to make a response is appropriate. Thus, simple reaction time is not a good proxy for driving ability on this task. Moreover, the Judgment task requires higher-order cognition associated with making and implementing a decision about moving obstacle avoidance. Such a task can also involve different strategies, where one driver may choose to take a more cautious approach even if it does not necessarily lead to a safer outcome, which could work against this

measure being predictive of on-road success.

#### 4.3. Short-term memory

On the Memory task, participants are shown a geometric shape and then asked to recreate it after a brief delay. Task duration and replication success were both good predictors of future on-road test outcomes: participants who correctly replicated more shapes and completed the task more quickly were more likely to pass their road test than participants who replicated fewer shapes and completed the task more slowly. Between-group comparisons also showed significant differences on both measures for light vehicle and truck drivers but not heavy vehicle drivers. This difference could be explained by the over-representation of return-to-work participants amongst the heavy vehicle group, however, as these participants scored unusually poorly on the Vitals despite being

**Table 7**

Mean and median performance on Vitals tasks by driving group (Heavy, Light, and Trucking) after excluding the return-to-work group (n = 2091).

Vehicle Type Variable	Heavy				Light				Trucking			
	N	Mean	SD	Median	N	Mean	SD	Median	N	Mean	SD	Median
Reaction Time Stage 2	347	0.511	0.061	0.512	1580	0.523	0.068	0.516	153	0.502	0.08	0.498
Judgement Premature Starts Stage 1	347	0.041	0.075	0.022	1585	0.045	0.08	0.022	154	0.045	0.11	0.019
Judgement Premature Starts Stage 2	347	0.058	0.059	0.024	1585	0.057	0.059	0.024	154	0.05	0.05	0.024
Judgement Success Stage 1	347	0.91	0.145	0.972	1584	0.892	0.161	0.971	153	0.88	0.2	0.971
Judgement Success Stage 2	347	0.773	0.113	0.791	1584	0.772	0.112	0.792	154	0.764	0.14	0.791
Judgement Reaction Time Stage 1	347	2.224	1.5	1.734	1584	2.45	1.51	1.989	153	2.261	1.548	1.788
Judgement Reaction Time Stage 2	347	2.599	0.956	2.313	1584	2.765	1.065	2.5	154	2.603	0.957	2.429
Judgement Obstacle Count Stage 1	347	0.95	0.635	0.853	1584	1.07	0.648	0.879	153	0.993	0.697	0.863
Judgement Start Count Stage 2	347	1.827	0.276	1.832	1584	1.867	0.299	1.833	154	1.891	0.413	1.834
Memory Duration	347	115.54	23.885	114.245	1585	118.528	24.12	115.671	154	117.806	24.671	115.164
Memory Correct Shapes	347	0.548	0.17	0.557	1585	0.538	0.164	0.537	154	0.509	0.165	0.53
Control Time Inside the Circle	347	0.437	0.108	0.439	1585	0.386	0.113	0.389	154	0.427	0.098	0.426
Control Surprise Obstacles Avoided	347	0.902	0.08	0.919	1585	0.915	0.075	0.937	154	0.881	0.099	0.902

**Table 8**

Negative (NPV) and positive predictive value (PPV) of the classifier model validation dataset (n = 2164).

TN	TP	FN	FP	NPV	PPV	ACC	N (No Indeterminate)
77.15 % (1330)	2.38 % (41)	17.0 % (293)	3.48 % (60)	0.8194	0.4059	0.7952	2164 (1724)

**Table 9**

One-way ANOVA results on variables (with non-linear age effects removed) between high and low risk groups in the training dataset (n = 1343).

Variable	F	SS	$\eta_p^2$	P
Intercept:0	17.9544	2.67E + 06	0.0132	0
Judgement:1	40.8936	2.48E + 02	0.0296	0
Judgement:2	26.8782	7.06E + 01	0.0196	0
Reaction:3	8.6865	2.75E + 02	0.0064	0.0033
Memory:4	131.8788	4.06E + 02	0.0895	0
Memory:5	107.713	3.27E + 02	0.0744	0
Judgement:6	47.9623	3.57E + 02	0.0345	0

**Table 10**

Demographics of the training dataset (n = 1343).

	Age (M ± SD)	N	Risk Level (Low/High)
Low-Risk Group	34.34 ± 17.23	955	0/955 (0.0 %)
High-Risk Group	42.66 ± 17.16	388	388/388 (100.0 %)
Full Sample	36.74 ± 17.61	1343	388/955 (28.9 %)
Medical1	42.63 ± 15.53	93	3/93 (3.2 %)
Research1	32.75 ± 5.18	99	0/99 (0.0 %)
Research2	30.40 ± 9.43	269	123/269 (45.7 %)
Medical2	74.21 ± 14.86	86	30/86 (34.9 %)
Research3	20.31 ± 3.84	318	0/318 (0.0 %)
Research4	44.97 ± 12.83	186	184/186 (98.9 %)
Research5	69.64 ± 11.67	45	6/45 (13.3 %)
Commercial1	39.98 ± 11.50	58	23/58 (39.7 %)
Commercial2	38.31 ± 7.23	13	2/13 (15.4 %)
Commercial3	38.65 ± 9.22	176	17/176 (9.7 %)

more likely to pass their on-road test than other participants. When the return-to-work participants were excluded from analysis, these statistical differences disappeared.

Memory is a well-established proxy for general cognitive functioning. Working memory involves the ability to flexibly coordinate other mental abilities, be they semantic, sensory, or motoric (Baddeley, 1992; D'Esposito and Postle, 2015). Thus, poor or declining working memory function can have broad impacts on an individual's ability to perform complex tasks. Memory impairments are a common symptom of MCI, stroke, or diabetes, and declining memory function is often the first observable symptom of neurodegenerative conditions such as Alzheimer's disease (Alzheimer's Association, 2022; Creavin et al., 2016;

Huntley and Howard, 2010). Our results indicate that short-term memory performance is predictive of impaired driving ability. Driving involves regular use of short-term, working memory functions, such as remembering the information posted on road signs or the locations of other road users. Consequently, a driver whose short-term, working memory performance is low or declining may be a strong candidate to recommend for an on-road test of their driving ability.

#### 4.4. Sensorimotor Control

The sensorimotor Control task requires participants to maneuver a virtual rolling ball inside a target circle by tilting the tablet left or right like a steering wheel, while avoiding any sudden obstacles that appear. According to our model, time spent inside the target circle predicted future on-road test outcomes, but obstacle avoidance did not. The Control task is a complex task which, like driving itself, requires a variety of perceptual, cognitive, and motor skills in order to perform (CCMTA, 2022). Participants must use visual pursuit to follow the target as it moves across the screen, perform fluid movements to maintain the ball within the moving target, and automatically orient towards and react quickly to sudden obstacles. Participants may pursue different goods-based strategies during the task, such as whether to maximize time spent in the target circle or to avoid obstacles, depending upon the relative value that they place on each task measure, and must also deal with simple action-based decisions, such as whether to turn left or right to avoid an obstacle (Wisniewski et al., 2020).

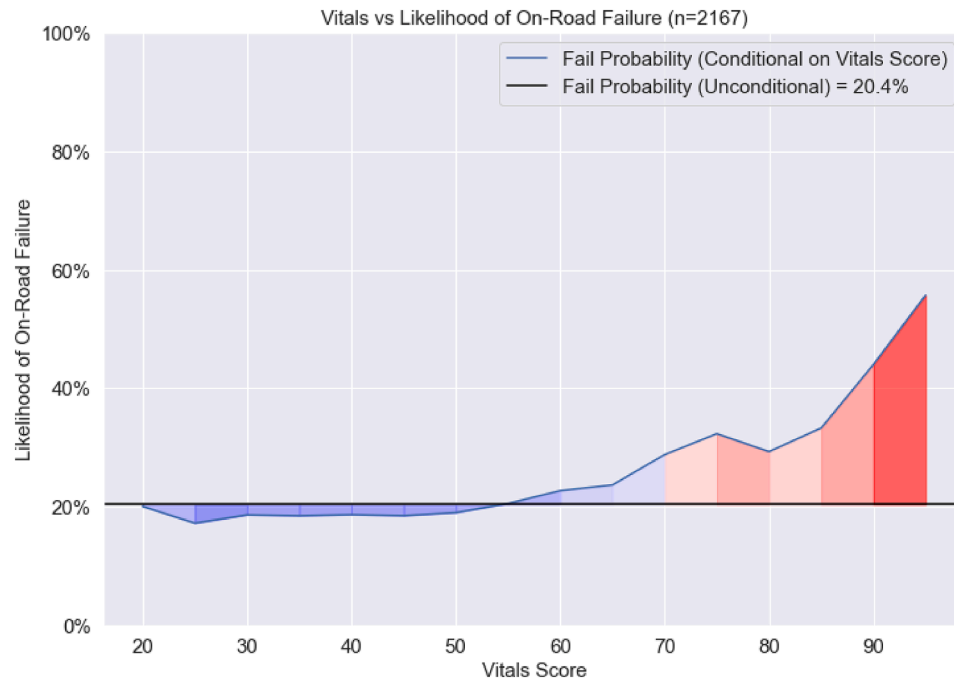
Given the complexity of the task, there may be no clear and simple explanation for why time in the circle but not obstacle avoidance predicts on-road test outcomes. Nevertheless, we can still offer some speculation and lay the groundwork for future study. Using fine motor movements to maintain one's vehicle within a visual lane is a critical driving skill that is continuously required of drivers. By contrast, sudden obstacle avoidance is still critical but much less frequently required, and is not deliberately tested during an on-road evaluation for obvious safety reasons. Failing to avoid an obstacle was much less critical during the Control task than it would be while operating a vehicle, and participants may have prioritized increased time on target over successful obstacle avoidance, thus leading to a dissociation between motor control skill and on-road performance for obstacle avoidance only. Therefore, performance at avoiding obstacles on the Control task may not relate to road test outcomes despite being theoretically relevant to real-world



**Table 11**

Accuracy comparison of model output for various data sources in the training dataset (n = 1343).

	ACC	FNR	FPR	TNR	TPR	Tri.Low	Tri.Up	Undefined
Full-sample	0.65	0.24	0.4	0.6	0.76	0.39	0.51	0.25
Medical1	0.89	0.67	0.1	0.9	0.33	0.39	0.51	0.07
Research1	0.84	–	0.16	0.84	–	0.39	0.51	0.36
Research2	0.47	0.49	0.57	0.43	0.51	0.39	0.51	0.31
Medical2	0.51	0.1	0.75	0.25	0.9	0.39	0.51	0.18
Research3	0.69	–	0.31	0.69	–	0.39	0.51	0.26
Research4	0.9	0.09	1	0	0.91	0.39	0.51	0.13
Research5	0.26	0.2	0.82	0.18	0.8	0.39	0.51	0.18
Commercial1	0.63	0.41	0.34	0.66	0.59	0.39	0.51	0.26
Commercial2	0.71	0.5	0.2	0.8	0.5	0.39	0.51	0.86
Commercial3	0.47	0.14	0.58	0.42	0.86	0.39	0.51	0.36

**Fig. 11.** Likelihood of on-road test failure depending on Vitals score in the validation dataset (n = 2167) using a smoothing width of 5 Vitals score points. Vitals score is calculated as the predicted probability of being impaired based on task performance. A higher score indicates worse performance.**Table 12**

Mixed Effect Logistic Model Results. The predictor variables in both models were selected based on the results of an ANOVA test, where they were found to be significant. The variance of the random effect term for Assessment Purpose in the Second Model was 0.2481 with a standard deviation of 0.498. One sample had an invalid Assessment Purpose label and was removed from the second model data.

Model	Dependent Variable	Predictor Variables	Link Function	Distribution	Residual Deviance	DOF	AIC	Significant Predictor Variables
Mixed Effect (Fixed Only)	Driving Outcome	Judgement Stage 2 Success, Judgement Stage 2 Go Count, Memory Shape Correctness, Control Time Inside Target	Logit	Binomial	5738.4	2145	5225.7	All predictor variables with p-values < 0.05
Mixed Effect (Fixed + Random Intercepts)	Driving Outcome	Same as First Model + Random Intercepts For Assessment Purpose	Logit	Binomial	5700.9	2144	5712.9	All predictor variables with p-values < 0.05

safe driving.

#### 4.5. Conclusion and limitations

In this study, we aimed to predict on-road driver performance in commercial heavy vehicle and truck drivers, as well as light vehicle drivers, depending on performance on the tablet-based Vitals cognitive screening tool. On-road driving evaluations are expensive, time-consuming, and can be unsafe for drivers who are most at-risk,

meaning that a reliable off-road screening tool can have substantial benefits for both drivers as well as the organizations or jurisdictions that administer the evaluation. The results of the study showed that performance on the Vitals was effective in predicting on-road performance across all three driving groups. Furthermore, drivers who failed their on-road evaluation performed significantly worse on several measures of the Vitals, including Judgement task success, correct shape replication on the Memory task, and time on-target in the Control task, compared to drivers who passed. A novel study of this nature necessarily has certain

**Table 13**

The fixed effect values for all predictor variables are consistent for all Assessment Purpose groups in Model 2. The difference in the intercept levels between Model 1 and Model 2 highlights the heterogeneity of the Assessment Purpose groups. However, the variation in the intercept levels is much smaller than the model intercept.

Assessment Purpose	Model 1 (Intercept)	Model 2 (Intercept)	Model Differences	Model 2 Random Effects (Deviation From Model 2 Baseline)
Other	3.6652976	2.862484	0.8028136	−0.21978381
Performance Evaluation	3.6652976	3.547325	0.1179676	0.46505697
Periodic Wellness	3.6652976	2.983414	0.6818836	−0.09885439
Post-Incident	3.6652976	2.774473	0.8908236	−0.30779521
Pre-hire	3.6652976	3.771767	−0.1062704	0.68949885
Evaluation				
Return to Work	3.6652976	2.588348	1.0769496	−0.49392004

limitations. Data sources for the classifier model had unequal sample sizes, and oversampling of small sources could not be used since some sample sizes were very small. Thus, all sources were weighted equally when building the model. Participant demographic data did not include information on gender, and precise vehicle models operated by each participant were also not collected (participants were sorted into groups based on the source of their data), so each vehicle category (Light, Heavy, and Trucking) contains various vehicle models which may differ in their utility and physical dimensions. Participants in the heavy vehicle group were disproportionately recruited to the study via return-to-work procedures after a safety incident at work, although these participants were excluded from a re-analysis of the univariate data without major impact on our conclusions. In sum, this is the first study to demonstrate the effectiveness of the Vitals tool among commercially licensed drivers. Whether drivers are seeking to renew their license or resume driving after a health or safety incident, the use of the Vitals as a screening tool can benefit both commercial vehicle drivers and evaluators in distinguishing between drivers who are likely to pass or fail an on-road evaluation.

### Role of the funding source

Funding for the conduct of this study and preparation of the article, including data collection, analysis, and writing, was provided by a National Sciences and Engineering Research Council (NSERC) Collaborative Research and Development grant awarded to Dr. Anthony Singhal and Impirica, and was also provided by Impirica.

### CRedit authorship contribution statement

**Daniel Scott:** Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Verification, Visualization, Writing – original draft, Writing – review & editing. **Alice Elizabeth Atkin:** Conceptualization, Visualization, Writing – original draft, Writing – review & editing. **Aaron Granley:** Conceptualization, Methodology, Software. **Anthony Singhal:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

### Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Anthony Singhal reports financial support was provided by Natural Sciences and Engineering Research Council of Canada. Aaron Granley reports financial support was provided by Impirica. Daniel Scott reports

financial support, administrative support, and equipment, and supplies were provided by Impirica. Daniel Scott reports a relationship with Impirica that includes: employment and equity or stocks. Aaron Granley reports a relationship with Impirica that includes: employment and equity or stocks.

### Data availability

The data that has been used is confidential.

### References

- Abu Dabrh, A.M., Firwana, B., Cowl, C.T., Steinkraus, L.W., Prokop, L.J., Murad, M.H., 2014. Health assessment of commercial drivers: a meta-narrative systemic review. *BMJ Open* 4 (3). <https://doi.org/10.1136/bmjopen-2013-003434>.
- Alzheimer's Association, 2022. Alzheimer's Disease Facts and Figures. <https://www.alz.org/media/documents/alzheimers-facts-and-figures.pdf>.
- Anderson, M.L., Auffhammer, M., 2014. Pounds that kill: the external costs of vehicle weight. *Rev. Econ. Stud.* 81 (2), 535–571. <https://doi.org/10.1093/restud/rdt035>.
- Baddeley, A., 1992. Working memory. *Science* 255 (5044), 556–559. <https://doi.org/10.1126/science.1736359>.
- Bakhtiar, R., Tomczak, M.V., Langor, S., Scanlon, J.E.M., Granley, A., Singhal, A., 2020. Application of tablet-based cognitive tasks to predict unsafe drivers in older adults. *Transp. Res. Interdisciplinary Perspect.* 4, 100105 <https://doi.org/10.1016/j.trip.2020.100105>.
- Balaban, C., Hoffer, M.E., Szczupak, M., Snapp, H., Crawford, J., Murphy, S., Kiderman, A., 2016. Oculomotor, vestibular, and reaction time tests in mild traumatic brain injury. *PLoS One* 11 (9), e0162168.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67 (1), 1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Berndt, A., May, E., Clark, M., 2007. Drivers with Dementia: Environment, Errors and Performance Outcomes. International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, Stevenson, Washington. [https://drivingassessment.uiowa.edu/sites/drivingassessment.uiowa.edu/files/wysiwyg\\_uploads/067\\_berndtclark.pdf](https://drivingassessment.uiowa.edu/sites/drivingassessment.uiowa.edu/files/wysiwyg_uploads/067_berndtclark.pdf).
- Bondallaz, P., Favrat, B., Chtioui, H., Fornari, E., Maeder, P., Giroud, C., 2016. Cannabis and its effects on driving skill. *Forensic Sci. Int.* 268, 92–102. <https://doi.org/10.1016/j.forsciint.2016.09.007>.
- Bureau of Labor Statistics (2021). National census of fatal occupation injuries in 2020. Retrieved from <https://www.bls.gov/iif/oshcfoi1.htm>.
- Canadian Council for Motor Transportation Administrators, 2022. National safety code for motor carriers. <https://www.ccmta.ca/en/national-safety-code>.
- Castillo-Manzano, J.I., Castro-Nuño, M., 2012. Driving licenses based on points systems: Efficient road safety strategy or latest fashion in global transport policy? A worldwide meta-analysis. *Transp. Policy* 21, 191–201. <https://doi.org/10.1016/j.tranpol.2012.02.003>.
- Choi, S.Y., Yoo, D.H., Lee, J.S., 2015. Usefulness of the driveABLE cognitive assessment in predicting the driving risk factor of stroke patients. *J. Phys. Ther. Sci.* 27(10), 3133–3135. <https://doi.org/10.1589/jpts.27.3133>.
- Clermont, Y. (2018). Impaired driving statistics: Telling Canada's story in numbers. Standing Senate Committee on Legal and Constitutional Affairs. Retrieved from [https://sencanada.ca/content/sen/committee/421/LCJC/Briefs/2018-LCJC\\_StatCan\\_e.pdf](https://sencanada.ca/content/sen/committee/421/LCJC/Briefs/2018-LCJC_StatCan_e.pdf).
- Compton, R.P., Berning, A., 2015. Drug and alcohol crash risk. <https://rosap.ntl.bts.gov/view/dot/1993>.
- Compton, R.P., 2017. Marijuana-impaired driving: A report to congress. Washington, D.C. <https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/812440-marijuana-impaired-driving-report-to-congress.pdf>.
- Creavin, S.T., Wisniewski, S., Noel-Storr, A.H., Trevelyan, C.M., Hampton, T., Rayment, D., Cullum, S., 2016. Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *Cochrane Database Syst. Rev.* 1, CD011145. <https://doi.org/10.1002/14651858.CD011145.pub2>.
- Dai, H., Bragazzi, N.L., Younis, A., Zhong, W., Liu, X., Wu, J., Grossman, E., 2021. Worldwide trends in prevalence, mortality, and disability-adjusted life years for hypertensive heart disease from 1990 to 2017. *Hypertension* 77 (4), 1223–1233. <https://doi.org/10.1161/HYPERTENSIONAHA.120.16483>.
- D'Esposito, M., Postle, B.R., 2015. The cognitive neuroscience of working memory. *Annu. Rev. Psychol.* 66, 115–142. <https://doi.org/10.1146/annurev-psych-010814-015031>.
- Dingus, T.A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., Hankey, J., 2016. Driver crash risk factors and prevalence evaluation using naturalistic driving data. *PNAS*, 113(10), 2636–2641. <https://doi.org/https://doi.org/10.1073/pnas.1513271113>.
- Dobbs, A.R., 1997. Evaluating the driving competence of dementia patients. *Alzheimer Dis. Assoc. Disord.* 11 (Suppl 1), 8–12. <https://doi.org/10.1097/00002093-199706001-00003>.
- Dobbs, A.R., Heller, R.B., Schopflocher, D., 1998. A comparative approach to identify unsafe older drivers. *Accid. Anal. Prev.* 30 (3), 363–370. [https://doi.org/10.1016/S0001-4575\(97\)00110-3](https://doi.org/10.1016/S0001-4575(97)00110-3).
- Dobbs, A.R., 2013. Accuracy of the DriveABLE cognitive assessment to determine cognitive fitness to drive. *Can. Fam. Physician*, 59(3), e156–161.

- Dumurgier, J., Tzourio, C., 2020. Epidemiology of neurological diseases in older adults. *Rev. Neurol.* 176 (9), 642–648. <https://doi.org/10.1016/j.neuro.2020.01.356>.
- Federal Motor Carrier Safety Administration (2007). Unit costs of medium and heavy truck crashes. Retrieved from: <https://www.fmcsa.dot.gov/sites/fmcsa.dot.gov/files/docs/UnitCostsTruck%20Crashes2007.pdf>.
- Federal Motor Carrier Safety Administration (2009). The bus crash causation study report to Congress. Retrieved from <https://www.fmcsa.dot.gov/mission/policy/bu-s-crash-causation-study-report-congress>.
- Fell, J.C., Scherer, M., 2017. Estimation of the Potential Effectiveness of Lowering the Blood Alcohol Concentration (BAC) Limit for Driving from 0.08 to 0.05 Grams per Deciliter in the United States. *Alcohol Clin. Exp. Res.* 41 (12), 2128–2139. <https://doi.org/10.1111/acer.13501>.
- Goldenfeld, C., Blom, M., Houwing, S., 2016. Serious alcohol offenders in traffic: Extent of the problem and characteristics of the offenders. Retrieved from <https://swov.nl/system/files/publication-downloads/r-2016-12.pdf>.
- Gorus, E., De Raedt, R., Lambert, M., Lempers, J.C., Mets, T., 2008. Reaction times and performance variability in normal aging, mild cognitive impairment, and Alzheimer's disease. *J. Geriatr. Psychiatry Neurol.* 21 (3), 204–218. <https://doi.org/10.1177/0891988708320973>.
- Government of Alberta, 2022a. Driver's licence suspensions and disqualifications. Retrieved October 6, 2022 from <https://www.alberta.ca/drivers-licence-suspensions-and-disqualifications.aspx#jumplinks-1>.
- Government of Alberta, 2022b. Information for aging drivers. Retrieved October 6, 2022 from <https://www.alberta.ca/information-for-aging-drivers.aspx>.
- Government of Alberta, 2022c. Report a medical condition. Retrieved October 6, 2022 from <https://www.alberta.ca/report-a-medical-condition.aspx>.
- Governor's Highway Safety Association, 2022. Spotlight on highway safety: pedestrian traffic fatalities by state. <https://www.ghsa.org/resources/Pedestrians22>.
- Hancock, P.A., Desmond, P.A. (Eds.), 2000. *Stress, Workload, and Fatigue*. Routledge.
- Hartman, R.L., Huestis, M.A., 2013. Cannabis effects on driving skills. *Clin. Chem.* 59 (3), 478–492. <https://doi.org/10.1373/clinchem.2012.194381>.
- Huntley, J.D., Howard, R.J., 2010. Working memory in early Alzheimer's disease: a neuropsychological review. *Int. J. Geriatr. Psychiatry* 25 (2), 121–132. <https://doi.org/10.1002/gps.2314>.
- Ichikawa, M., Nakahara, S., Inada, H., 2015. Impact of mandating a driving lesson for older drivers at license renewal in Japan. *Accid. Anal. Prev.* 75, 55–60. <https://doi.org/10.1016/j.aap.2014.11.015>.
- Impirica, 2022. DriveABLE Cognitive Assessment Tool (DCAT). Retrieved October 7, 2022 from <https://impirica.tech/solutions/dcat/>.
- International Transport Forum, 2020. Road safety annual report 2020. <https://www.itf-oecd.org/road-safety-annual-report-2020>.
- International Transport Forum, 2021. Road safety data: Canada. <https://www.itf-oecd.org/road-safety-annual-report-2021-impact-covid-19>.
- Johansson, R., 2009. Vision zero: implementing a policy for traffic safety. *Saf. Sci.* 47, 826–831. <https://doi.org/10.1016/j.ssci.2008.10.023>.
- Kessels, R.P., van Zandvoort, M.J., Postma, A., Kappelle, L.J., de Haan, E.H., 2000. The Corsi Block-Tapping Task: standardization and normative data. *Appl. Neuropsychol.* 7 (4), 252–258. [https://doi.org/10.1207/S15324826AN0704\\_8](https://doi.org/10.1207/S15324826AN0704_8).
- Kim, E., Muennig, P., Rosen, Z., 2017. Vision zero: a toolkit for road safety in the modern era. *Inj. Epidemiol.* 4 (1), 1. <https://doi.org/10.1186/s40621-016-0098-z>.
- Korner-Bitensky, N., Sofer, S., 2009. The DriveABLE Competence Screen as a predictor of on-road driving in a clinical sample. *Aust. Occup. Ther. J.* 56 (3), 200–205. <https://doi.org/10.1111/j.1440-1630.2008.00749.x>.
- Kowalski, K., Tuokko, H., 2007. On-road driving assessment of older adults: A review of the literature. <https://core.ac.uk/download/pdf/236984963.pdf>.
- Langford, J., Fitzharris, M., Koppel, S., Newstead, S., 2004. Effectiveness of mandatory license testing for older drivers in reducing crash risk among urban older Australian drivers. *Traffic Inj. Prev.* 5 (4), 326–335. <https://doi.org/10.1080/15389580490509464>.
- Li, S., 2012. Traffic safety and vehicle choice: quantifying the effects of the 'arms race' on American roads. *J. Appl. Economet.* 27, 34–62. <https://doi.org/10.1002/jae.1161>.
- Matsui, Y., Oikawa, S., Sorimachi, K., Imanishi, A., Fujimura, T., 2016. Association of Impact Velocity with Risks of Serious Injuries and Fatalities to Pedestrians in Commercial Truck-Pedestrian Accidents. *Stapp Car Crash J.* 60, 165–182. <https://doi.org/10.4271/2016-22-0007>.
- Matthews, F.E., Stephan, B.C.M., Robinson, L., Jagger, C., Barnes, L.E., Arthur, A., Brayne, C., 2016. A two decade dementia incidence comparison from the Cognitive Function and Ageing Studies I and II. *Nat. Commun.* 7, 11398. <https://doi.org/10.1038/ncomms11398>.
- Mitchell, C.G., 2008. The licensing of older drivers in Europe - A case study. *Traffic Inj. Prev.* 9 (4), 360–366. <https://doi.org/10.1080/15389580801895160>.
- Monsell, S., 2003. Task switching. *Trends Cogn. Sci.* 7 (3), 134–140. [https://doi.org/10.1016/s1364-6613\(03\)00028-7](https://doi.org/10.1016/s1364-6613(03)00028-7).
- Organization for Economic Co-operation and Development, 2022. Road accidents (indicator). <https://data.oecd.org/transport/road-accidents.htm>.
- Pachella, R.G., 1974. The interpretation of reaction time in information processing research. Routledge.
- Robbins, T.W., 2007. Shifting and stopping: fronto-striatal substrates, neurochemical modulation and clinical implications. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 362 (1481), 917–932. <https://doi.org/10.1098/rstb.2007.2097>.
- Robertson, D.A., Savva, G.M., Coen, R.F., Kenny, R.A., 2014. Cognitive function in the prefrailty and frailty syndrome. *J. Am. Geriatr. Soc.* 62 (11), 2118–2124. <https://doi.org/10.1111/jgs.13111>.
- Rock, S.M., 1998. Impact from changes in Illinois drivers license renewal requirements for older drivers. *Accid. Anal. Prev.* 30 (1), 69–74. [https://doi.org/10.1016/s0001-4575\(97\)00063-8](https://doi.org/10.1016/s0001-4575(97)00063-8).
- Savage, I., 2013. Comparing the fatality risks in United States transportation across modes and time. *Res. Transp. Econ.* 43, 9–22. <https://doi.org/10.1016/j.retrec.2012.12.011>.
- Saxby, D.J., Matthews, G., Warm, J.S., Hitchcock, E.M., Neubauer, C., 2013. Active and passive fatigue in simulated driving: discriminating styles of workload regulation and their safety impacts. *J. Exp. Psychol. Appl.* 19 (4), 287–300. <https://doi.org/10.1037/a0034386>.
- Schmitt, A., 2020. *Right of way: Race, class, and the silent epidemic of pedestrian deaths in America*. Island Press.
- Sharma, M., Nazareth, I., Petersen, I., 2016. Trends in incidence, prevalence and prescribing in type 2 diabetes mellitus between 2000 and 2013 in primary care: a retrospective cohort study. *BMJ Open*, e010210. <https://doi.org/doi:10.1136/bmjopen-2015-010210>.
- Silva, K.G., Aloise, D., Xavier-de-Sousa, S., Mladenovic, N., 2018. Less is more: simplified Nelder-Mead method for large unconstrained optimization. *Yugosl. J. Oper. Res.* 28 (2), 153–169. <https://doi.org/10.2298/YJOR180120014G>.
- Singh, S., 2015. Critical reasons for crashes investigated in the national motor vehicle crash causation survey. (Traffic Safety Facts - Crash Stats, Issue. N.H.T.S.A. Analysis). <http://www.nrd.nhtsa.dot.gov/Pubs/812115.pdf>.
- Statistics Canada, 2021. Population estimates on July 1st, by age and sex. Retrieved from <https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000501&pickMembers%5B0%5D=1.1&pickMembers%5B1%5D=2.1&cubeTimeFrame.startYear=1971&cubeTimeFrame.endYear=2021&referencePeriods=19710101%2C20210101>.
- Stuss, D.T., Stethem, L.L., Hugenholtz, H., Picton, T., Pivik, J., Richard, M.T., 1989. Reaction time after head injury: fatigue, divided and focused attention, and consistency of performance. *J. Neurol. Neurosurg. Psychiatry* 52 (6), 742–748. <https://doi.org/10.1136/jnnp.52.6.742>.
- Teichner, W.H., 1954. Recent studies of simple reaction time. *Psychol. Bull.*, 51(2:1), 128–149. <https://doi.org/10.1037/h0060900>.
- Tomczak, M.V., Bakhtiari, R., Langor, S. D., Granley, A., Campbell, F., Singhal, A., 2022. The effects of cannabis and cocaine on driving related tasks of perception, cognition, & action. Unpublished results.
- Treisman, A., 1991. Search, similarity, and integration of features between and within dimensions. *J. Exp. Psychol. Hum. Percept Perform* 17 (3), 652–676. <https://doi.org/10.1037/0096-1523.17.3.652>.
- Vanlaar, W., Hing, M.M., Robertson, R., Mayhew, D., Carr, D., 2016. An evaluation of Ontario's Group Education Session (GES) for license renewal of seniors aged 80 and above. *J. Safety Res.* 56, 33–39. <https://doi.org/10.1016/j.jsr.2015.11.009>.
- Wispinski, N.J., Gallivan, J.P., Chapman, C.S., 2020. Models, movements, and minds: bridging the gap between decision making and action. *Ann. N Y Acad. Sci.* 1464 (1), 30–51. <https://doi.org/10.1111/nyas.13973>.
- World Health Organization, 2018. Summary: global status report on road safety. <http://www.who.int/publications/i/item/9789241565684>.
- World Health Organization, 2021. Global Plan: Decade of Action for Road Safety 2021–2030. <https://www.who.int/publications/m/item/global-plan-for-the-decade-of-action-for-road-safety-2021-2030>.
- Zou, W., Wang, X., Zhang, D., 2017. Truck crash severity in New York City: An investigation of the spatial and the time of day effects. *Accid. Anal. Prev.* 99, 249–261. <https://doi.org/10.1016/j.aap.2016.11.024>.