Driving Tradeoffs with a Touchscreen: Workload Measurements with a Detection and Choice Response Task

Vehicle-based technologies that compete for attention have become ubiquitous in modern vehicles, and should be assessed for their visuo-manual and cognitive workload. The current International Standards Organization (ISO) Detection Response Task (DRT) is sensitive to overall mental workload, but the current standard advises against utilizing the DRT for tasks with manual interference. However, this recommendation has not been thoroughly investigated. We examined the potential for the ISO DRT and a Choice Response Task (CRT) to: (1) detect the magnitude of cognitive and visuo-manual workload, (2) demonstrate tradeoffs, and (3) produce manual interference for a steering task and varying difficulty of a touchscreen visual search task. The data demonstrate DRT and CRT sensitivity to the presence but not difficulty of visual search on a touch screen, while the CRT demonstrates tradeoffs with steering and visual search, possibly due to visuo-manual interference. This investigation furthers our understanding of measures for workload components.

INTRODUCTION

Our limited capabilities to visually perceive a dynamic environment, decide quickly upon actions, and manually execute these actions determine the success with which we operate motor vehicles. However, previous researchers pinpoint diversions of attention to secondary, non-driving related tasks as the leading factor of driver crashes and near crashes (Klauer et al., 2014). Regan, Hallett, and Gordon, (2011) define driver distraction as "the diversion of attention away from activities critical for safe driving toward a competing activity." In-vehicle information system (IVIS) technology, such as touchscreen interfaces, allow for an increase in secondary tasks while driving, but have the potential to divert valuable limited resources away from critical safe-driving activities. These diversions, or distractions, incorporate a combination of visual and cognitive components, which are highly correlated but distinctly separable psychological constructs (Strayer, Watson, & Drews, 2011).

Strayer, Watson, and Drews (2011) established a theoretical framework for classifying aspects of driver distraction into visual, manual, and cognitive workload. Researchers utilize a myriad of techniques to measure these components of workload in the lab, but many of these methods become unreliable or difficult to implement when utilized for naturalistic driving. In order to address measuring one component, the International Standards Organization (ISO) Detection Response Task (DRT) (see ISO 17488, 2015) was created to provide a feasible and reliable measure of cognitive workload in the vehicle. The device consists of a button attached to the participant's thumb or finger, and a visual or haptic stimulus. To assess cognitive workload, the DRT allows for a millisecond-accurate button response to a light or vibration, which occurs every 3-5 seconds. The ISO DRT has successfully detected the cognitive effects of secondary tasks while driving in multiple studies (e.g., Conti et al, 2012; Mantzke & Keinath, 2015)

However, the capacity of the DRT to discriminate between visuo-manual and cognitive distraction has only recently been investigated. Two previous studies have assessed different strategies for simultaneously detecting and discriminating between visual and cognitive workload with a novel configuration of head-mounted (HM) lights, vibro-tactile (VT) devices, and dash-mounted (DM) lights (Castro, Cooper, & Strayer, 2016; Cooper, Castro, & Strayer, 2016). Findings from these studies demonstrate that DRT measurements can be differentially sensitive to cognitive and visual workloads, but don't necessarily provide the magnitude of distraction proportionally assigned to the different types of workload.

Considering that previous work has found differential effects for the three components of workload, empirical support exists for creating recommendations for each of these components. However, the National Highway Traffic Safety Administration (NHTSA, 2013) establishes guidelines for detecting only manual and visual distractions. Further work demonstrates that the cognitive demands of tasks during driving interact with visual and manual workload (Recarte, Pérez, Conchillo, & Nunes, 2008). Both visual and cognitive demands contribute to overall driver distraction, but researchers have historically focused upon one type of distraction at a time, or simply measured overall demand (see Harms, 1991; Strayer, Turrill, Coleman, & Cooper, 2014).

The aim of the current study is to test novel measurements in their ability to differentially detect the magnitude of workload in a task that involves all three components (visual, manual, and cognitive). In this study we tested the ability of an ISO DRT and a Choice Response Task (CRT) to detect the impact of a visual search task on lateral steering deviation in a driving simulator. In the first configuration (Experiment 1), participants responded to an ISO standard flashing red light mounted on the dashboard. In the second configuration (Experiment 2), participants responded to a modified DM light that produces two intensities of light. Both experiments utilized different difficulties of a visual search task (Treisman & Gelade, 1980) and a baseline for comparison.

We hypothesized slower reaction times (RT) and lower hit rates under visual loads to the dash mounted DRT (Experiment 1) proportional to the difficulty of the visual search task. We also expected a tradeoff in performance between steering deviation and the difficulty of the visual search task, but little to no tradeoff between DRT performance and the other tasks' performance. Finally, we hypothesized little to no manual interference between the DRT and the visual search task performance. For the CRT (Experiment 2), we predicted poorer detection for the difficulty of the visual search task (see Castro, Strayer, Matzke, & Heathcote, Under Review), and tradeoffs between the CRT and the other tasks. We also expected to find evidence of manual interference for the right-hand response, as both the visual search task and the right-handed CRT responses shared a manual response.

METHOD

Participants

Participants were recruited from a pool of undergraduate students enrolled in a psychology course at the University of Utah. In Experiment 1, thirty-seven participants 18 to 36 years old (M = 22.5) completed the study. In Experiment 2, twenty-three participants 18 to 41 years old (M = 23.6) completed the study. Participants received course credit as compensation upon completion of the one-hour study.

Materials

The background task was displayed upon a 106 cm diagonal Samsung Television monitor (1920 x 1080 pixels). Participants sat in an L3 communications MPRI Ship Analytics, Simulation Technology Solutions Simulator, manufactured by I-Sim Corporation. However, in line with Castro, Cooper, and Strayer (2016) the simulator only provided the background screen, dashboard, steering wheel, and seat. The screens did not display a realistic driving simulation. Instead, the forward screen displayed a lateral steering tracking task described below. A Toshiba P55T-A5202 Touchscreen computer (39.62 cm diagonal) was fixed approximately 28 degrees down and 28.7 degrees to the right of the tracking task target location. A rotary encoder attached to the steering wheel of the driving simulator recorded deviation from the target in the background task. The visual search task was selected from the PEBL psychology experiment building language battery (Mueller & Piper, 2014).

Tasks. Both experiments included the three following tasks: The tracking task, the visual search task, and the DRT or CRT. The tracking task was displayed on the simulator television monitor. Participants used a steering wheel to control a triangle. Above the triangle, a yellow ball moved in a pseudo-random pattern. Participants were instructed to keep the triangle pointed at the ball in a typical tracking fashion.

For the visual search task, participants were asked to respond by touching either a green or white "O" amongst 10 white "C" and "o" distractors with the index finger of their right hand. The visual search task contained either a target or no target with 10 distractors. There was no feedback for correct or incorrect selection of targets. After selection, a new set of distractors and a target were immediately displayed.

The participants responded to either a DM DRT or a DM CRT in order to assess workload components. Both devices presented a stimulus upon the dash of the driving stimulator. The DRT device followed the specifications outlined in ISO 17488 (ISO, 2015). For the CRT device, an LED light presented two intensities of light, and the participant responded to the brighter light with a button attached to one thumb and the lower-intensity light with a button attached to the other thumb. These responses were counterbalanced within a participant.

Procedure

Participants in Experiment 1 were assigned to use the ISO DM DRT for the duration of the experiment. In Experiment 2, participants were assigned to use the DM CRT. Both experiments were conducted across 16 counterbalanced blocks for one minute, which all included the tracking task and responding to the response tasks. These tasks were considered the baseline in both experiments. Participants also completed three dual task blocks with a visual search task (Treisman & Gelade, 1980) on the touchscreen, which included a green target (easy), a white target (difficult), or a mix of green and white trials (mixed). These blocks were counterbalanced using a balanced Latin Square Design. Researchers instructed participants on the task and trained participants on each task for approximately one minute. Participants also received a break between blocks.

In Experiment 1, participants responded to an LED light that flashed red for one second, mounted on the dashboard, approximately 61 cm from the participants. The DRT randomly presented a light every 3-5 seconds. The lights remained illuminated for one second or until a response was made. In Experiment 2, participants responded to an LED light in the same location, but two intensities of light were presented. Participants responded to the bright and dim lights with buttons attached to both thumbs. In Experiment 1, participants responded by pressing the button attached to only their left hand against the steering wheel, but in Experiment 2, participants used both hands.

Measures. For Experiment 1, RTs to the presentation of a light were recorded as well as number of hits and misses with regard to detection of the stimulus. Hits constituted responses that occurred 100-2500 milliseconds after the onset of a stimulus. Non-responses and responses outside of this window were recorded as a miss. For Experiment 2, RT was recorded as well as the proportion of correct responses for each hand. Hits and misses were also recorded. Steering error was calculated through deviation of the triangle from the ball using the rotary encoder. Previous studies using eye tracking measurements for the forward screen demonstrated that DRT measures of visual load accurately detected the proportion of time eyes were on the forward roadway, which correlated with steering error (Castro, Cooper, & Strayer, 2016; Cooper, Castro, & Strayer, 2016). Therefore, we can assume a strong relationship between the percent of time the eyes were directed to the forward roadway and the percentage of hits to the DM DRT.

RESULTS

For the results, the open-source language and statistical computing environment R (R Core Team, 2016) was utilized to perform repeated measures ANOVAs and produce plots of the tested measures. All error bars represent 95% confidence intervals of the mean of subject means.

Experiment 1 – Visual Search Task

Reaction Time. RTs are reported in milliseconds. Repeated measures ANOVA found statistically significant differences in RT between easy (M = 1505, SD = 953), mixed (M = 1974, SD = 1094), and difficult search tasks, (M = 2418, SD = 1043) $F(2, 35) = 139.58$, $p < .01$, $\eta^2 = .98$, which demonstrated that visual task difficulty affected RT. (See Figure 1).

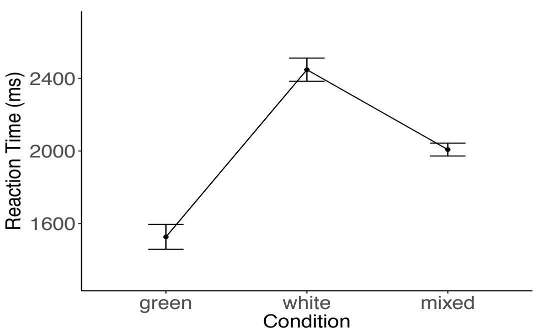

Figure 1. Response time in milliseconds to the easy (green), difficult (white), and mixed (both target colors) conditions for the visual search task. Visual search took longer when the distractors shared the same color as the target.

Accuracy. A repeated measures ANOVA did not find a significant difference between easy (M = 99.54%, SD = .79%) difficult (M = 99.71%, SD = .70%) and mixed (M = 99.35%, SD = .60%) conditions $F(2, 35)$ = , $p = .06$, $\eta^2 = .99$, demonstrating that task difficulty did not significantly change accuracy between the conditions for the visual search task.

Experiment 1 – DRT

Reaction Time. RTs are reported in milliseconds. Repeated measures ANOVA found statistically significant differences in RT between baseline (M = 372, SD = 88), easy (M = 576, SD = 67), mixed (M = 579, SD = 51), and difficult (M = 596, SD = 59) DRT responses $F(3, 35) = 55.19$, $p < .01$, $\eta^2 = .96$, which demonstrated that visual task difficulty affected DRT RT (See Figure 2). However, the magnitude of difficulty in the visual search task did not significantly affect the DRT RT in Tukey HSD post-hoc pairwise comparisons. Only baseline differed from the others.

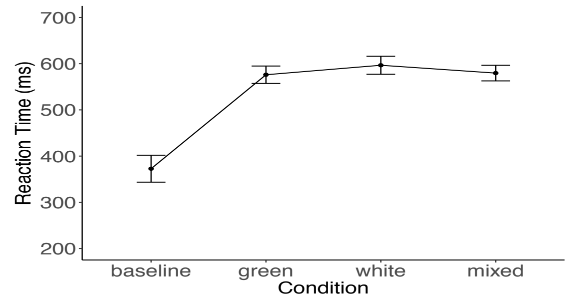

Figure 2. Reaction time in milliseconds to the DRT demonstrated a significant change from baseline to the presence of the visual search task. However, the difficulty levels of the visual search task did not significantly differ.

Hit Rate. The proportion of hits (classified as responses between 150 and 2500 milliseconds) and misses (after 2500 ms or no response) were compared with repeated-measures ANOVA, which indicated a significant effect of condition upon DRT hit rate $F(3, 35) = 12.4$, $p < .01$, $\eta^2 = .99$. However, post-hoc pairwise comparisons revealed that the only significant difference existed for the baseline condition (M = 98%, SD = 07%) compared to the others (M = 90%, SD = 06%); (M = 92%, SD = 05%); (M = 91%, SD = 05%) $p < .01$. This indicates that although the DRT is sensitive to the existence of a visuo-manual task, it is not sensitive to the differences in difficulty of that task. This finding confirms Castro, Cooper, and Strayer's (2016) conclusion that the red DM DRT did not significantly differentiate visual and cognitive workload, but it did successfully differentiate between levels of cognitive workload alone.

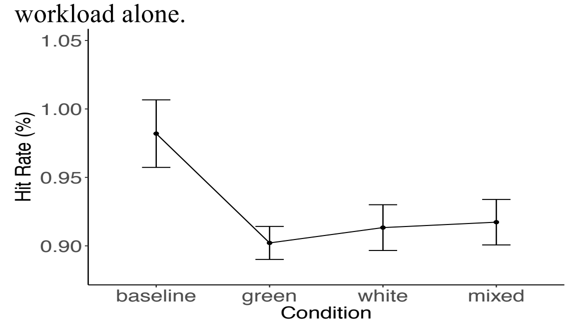

Figure 3. Hit rate as a percentage of successful responses compared to the total presentation of stimuli.

Experiment 1 – Steering Deviation

Root Mean Squared Error (RMSE) significantly differed across the four conditions $F(3,35) = 70.07$, $p < .01$ η^2 = . 99 according to repeated measures ANOVA (see Figure 4).

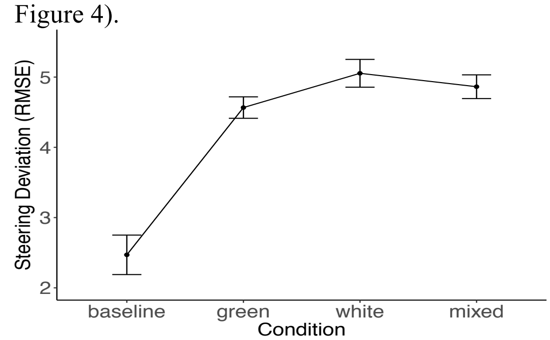

Figure 4. Deviation from tracking the optimal path of a ball across driving alone, easy, difficult, and mixed trial conditions.

Summary. Experiment 1 tested the sensitivity of a DRT to visuo-manual workload. The ISO standard light was predicted to be sensitive to the magnitude of difficulty in the visual search task. Despite our prediction, the DM DRT only appeared to detect the presence of a secondary task, and was not differentially sensitive to the levels of difficulty. A Pearson product-moment correlation coefficient was computed to assess the relationship between steering deviation and DRT RT, finding a strong correlation $r(35) = .66$, $p < .01$. However, there was not a significant correlation between the visual search task RT and steering deviation $r(35) = .10$, $p = .29$ or DRT RT $r(35) = .04$, $p = .67$. These data suggest that in Experiment 1, data from the DRT were unable to differentiate between visuo-manual levels of difficulty, but still maintained sensitivity toward performance outcomes in the form of steering error (as illustrated by the similarities between Figures 2 and 4). In order to determine the cause of these limitations, Experiment 2 utilized a CRT in order to test for potential cognitive tradeoffs and check for manual interference. In addition, the number of trials completed in the visual search task was calculated to determine other potential tradeoffs that might depend upon participant strategy.

Experiment 2 – Visual Search Task

Trials Completed. We found that the number of trials completed significantly differed between conditions according to repeated measures ANOVA $F(2,21) = 17.46$, $p < .01$, $\eta^2 = .77$. This measure may help to understand why accuracy did not significantly change by condition in Experiment 1 for the visual search task.

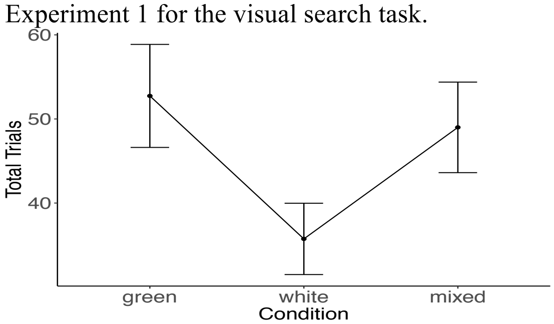

Figure 5. Average total number of trials completed by condition of the visual search task. Participants completed significantly fewer difficult (white) trials when the target matched the distractors.

Reaction Time. As in Experiment 1, repeated measures ANOVAs demonstrated that visual task responses were affected by condition, $F(2, 21) = 6.94$, $p < .01$, $\eta^2 = .96$. (see Figure 6).

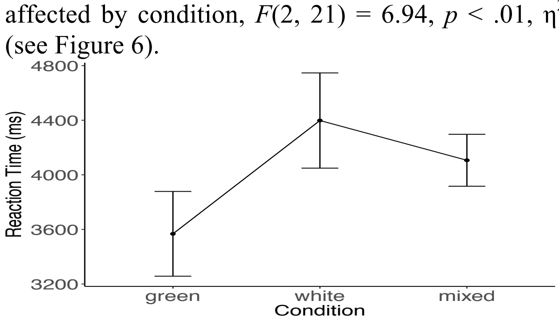

Figure 6. Response time in milliseconds to the visual search task. Responses were compared for the easy (green), difficult (white), and mixed (both target colors) conditions for the visual search task. Visual search took longer when the distractors shared the same color as the target.

Accuracy. In Experiment 2, accuracy was significantly affected by visual search conditions $F(2, 21) = 4.14$, $p = .03$ .01, $\eta^2 = .45$ (see Figure 7).

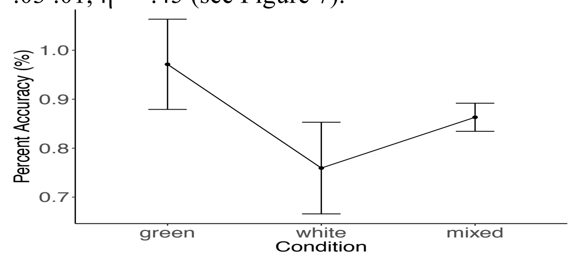

Figure 7. Percent accuracy of the visual search task significantly differed when the task had a green target, a white target, or a mix of targets. All distractors were white.

Experiment 2 – CRT

Reaction Time. RT was compared in repeated measures ANOVA for the correct trials of the CRT and was found to be significantly affected by condition $F(3,21) = 16.34$, $p < .01$, $\eta^2 = .76$. However, only baseline ($M = 611$, $SD = 74$) differed significantly from the others ($M = 742$, $SD = 37$); ($M = 764$, $SD = 53$); ($M = 717$, $SD = 66$) in pairwise comparisons $p < .05$.

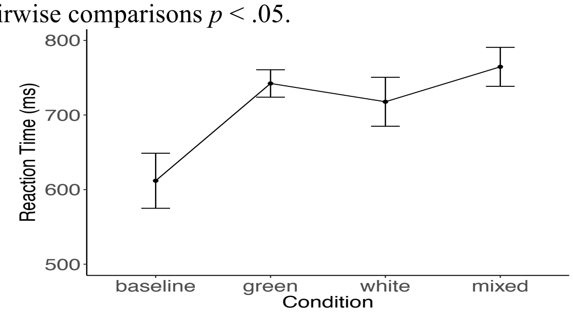

Figure 8. Reaction time in milliseconds to CRT correct trials demonstrated a significant change from baseline to the presence of the visual search task. However, the difficulty levels of the visual search task did not significantly differ.

Hit Rate and Accuracy. Hit rate and accuracy for the CRT were not significantly affected by condition in repeated measures ANOVA. We also did not find a significant difference between left and right hand responses. This provides indirect evidence that manual interference may not be the driving cause of the inability of the DRT or CRT to detect changes in visual difficulty.

Experiment 2 – Steering Deviation

In a repeated measures ANOVA, optimal tracking in the tracking task significantly differed across the four conditions, $F(3,21) = 49.07$, $p < .05$ $\eta^2 = .91$.

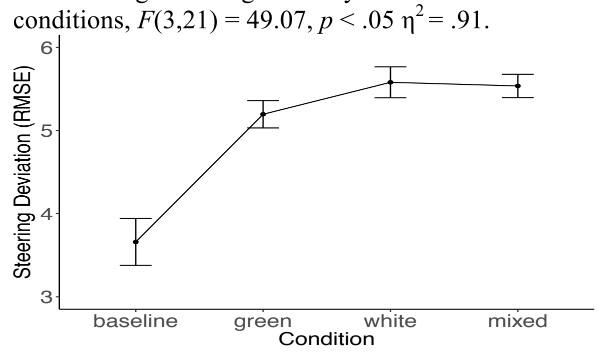

Figure 9. Deviation from tracking the optimal path of a ball across driving alone, easy, difficult, and mixed visual search conditions.

Experiments 1 and 2 – Comparisons

DRT and CRT tasks could detect the decrease in performance from baseline driving to performing the visual search task, but neither configuration was capable of finding differences in the difficulty of the visual search. In terms of tradeoff, both DRT and CRT RT correlated strongly with changes in steering deviation, with the CRT having a slightly weaker correlation $r(21) = .46$, $p < .01$. However, the CRT and the visual search task RTs were negatively correlated $r(21) = -.30$, $p < .01$, suggesting a tradeoff for the tasks.

DISCUSSION

In this study, we investigated the potential for the ISO DRT and a modified CRT to detect visuo-manual workload and manual interference. The ISO cautions against using the DRT for heavily manual tasks. In order to test this recommendation, we asked participants to perform a visual search task on a tablet while steering.

In Experiment 1, the DRT detected visual workload differences between baseline driving and performing the visual search task in RT. However, the DRT did not differentiate (in RT or hit rate) the levels of difficulty in the visual search task. This outcome supports the ISO recommendation, but does not directly find evidence of manual interference or tradeoffs between the DRT, driving-critical tasks, and non-critical secondary tasks.

In Experiment 2, the CRT also detected a correlation between steering performance and RT, but not hit rate or accuracy. Similar to the DRT, the CRT could not detect changes across the levels of difficulty for the visual search task. This finding suggests that the CRT does not improve detection of visuo-manual workload over the DRT. In addition, the CRT RT negatively correlated with the visual search task RT, which suggests a tradeoff of resources while attempting to attend to the two tasks simultaneously. However, no differences were found between the hands used to make responses, suggesting that this tradeoff was not due to manual interference. The patterns of mean deviation in the tracking task and their correlation with CRT RT suggest that RT would be the most sensitive dependent measure for detecting changes in visual workload. Despite the ease of implementation, the findings of this study suggest that the DRT and CRT should only be utilized to measure cognitive workload in conjunction with other measures for visual workload and manual interference.

As for the DRT affecting the driving environment, previous studies have found small, but significant effects of comparing baseline steering to steering with the DRT (see Castro, Strayer, Matzke, & Heathcote, Under Review). However, the DRT did not show a tradeoff with the tasks of interest, and the more intrusive CRT did demonstrate a significant effect in RT. Therefore, the DRT still appears to be a minimally invasive, easily implemented task for assessing cognitive workload in the vehicle. However, researchers should be aware of its limitations with regard to its sensitivity toward certain types of tasks.

REFERENCES

Castro, S.C., Strayer, D.L., Matzke, D., & Heathcote, A. (Under Review). Cognitive workload measurement and modeling under divided attention. *Journal of Experimental Psychology: General*.

Castro, S., Cooper, J., & Strayer, D. (2016, September). Validating Two Assessment Strategies for Visual and Cognitive Load in a Simulated Driving Task. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 60, No. 1, pp. 1899-1903). Sage CA: Los Angeles, CA: SAGE Publications.

Conti, A., Dlugosch, C., Vilimek, R., Keinath, A., & Bengler, K. (2012). An assessment of cognitive workload using detection response tasks. Advances in human factors and ergonomics series. Advances in human aspects of road and rail transportation, 735-743.

Cooper, J. M., Castro, S. C., & Strayer, D. L. (2016, September). Extending the Detection Response Task to Simultaneously Measure Cognitive and Visual Task Demands. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 60, No. 1, pp. 1962-1966). Sage CA: Los Angeles, CA: SAGE Publications.

Harms, L. (1991). Variation in drivers' cognitive load. Effects of driving through village areas and rural junctions. *Ergonomics, 34*(2), 151-160.

International Standards Organization. (2015). Road vehicles: Transport information and control systems. Detection- Response Task (DRT) for assessing selective attention in driving (ISO TC 22 SC 13 N17488). Working draft under development by Working Group 8 of ISO TC22, SC 13.

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. New England journal of medicine, 370(1), 54-59.

Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of neuroscience methods*, *222*, 250-259.

Mantzke, Odette, and Andreas Keinath. "Relating the detection response task to critical events–consequences of high cognitive workload to brake reaction times." Procedia Manufacturing 3 (2015): 2381-2386.

National Highway Traffic Safety Administration. (2013). Visual-manual NHTSA driver distraction guidelines for in-vehicle electronic devices. *Washington, DC: National Highway Traffic Safety Administration (NHTSA), Department of Transportation (DOT)*.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Recarte, M. Á., Pérez, E., Conchillo, Á., & Nunes, L. M. (2008). Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *The Spanish journal of psychology*, *11*(2), 374-385.

Regan, M. A., Hallett, C., & Gordon, C. P. (2011). Driver distraction and driver inattention: Definition, relationship and taxonomy. *Accident Analysis & Prevention*, *43*(5), 1771-1781.

Strayer, D.L., Turrill, J., Coleman, J., & Cooper, J.M. (2014). Measuring Cognitive Distraction in the Automobile II: Assessing In-Vehicle Voice-Based Interactive Technologies. AAA Foundation for Traffic Safety, Washington, DC.

Strayer, D. L., Watson, J. M., & Drews, F. A. (2011). Cognitive distraction while multitasking in the automobile. In Psychology of Learning and Motivation (Vol. 54, pp. 29-58). Academic Press.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. Cognitive psychology, 12(1), 97-136.