Validating Two Assessment Strategies for Visual and Cognitive Load in a Simulated Driving Task

Spencer Castro, University of Utah, Joel Cooper, University of Utah, David Strayer, University of Utah

With the emergence of vehicle-based technologies that could compete for attention due to visual and cognitive workloads in a driving environment, it is important to accurately assess the various components of potential distractions. Current Detection Response Task (DRT) measurements are sensitive to overall mental workload, but may not be useful for assessing visual workload. This study seeks to examine the ability of two unique extensions of DRTs to assess levels of cognitive and visual load in a lateral steering tracking task. Each DRT was tested in conditions that manipulated cognitive load, visual load, the combination of cognitive and visual load, and normal driving conditions. The data suggest that an altered design of the DRT may allow for reliable assessment of cognitive and visual loads simultaneously during a driving task. Measuring the components of different types of workload that lead to driver distraction may inform industry standards for assessing driver distraction in the vehicle.

INTRODUCTION

Successful driving performance requires a combination of visual and cognitive attention. Because visual attention is closely correlated with the physical location of the fovea, the National Highway Traffic Safety Administration (NHTSA, 2012) has established clear guidelines for the measurement of visual demand during driving tasks. However, cognitive demands of tasks during driving are another important aspect of driver distraction, but are not mentioned in the standard (see Harms, 1991; Strayer, Turrill, Coleman, & Cooper, 2014). Both visual and cognitive demands contribute to overall driver distraction, but current Detection Response Task (DRT) measurements do not separately account for visual and cognitive workloads. By varying the configuration and properties of the DRT, this study aims to provide evidence for a new DRT set which is simultaneously sensitive to visual and cognitive workloads in a distracted driving scenario.

Regan, Hallett, and Gordon, (2011) defined driver distraction as “the diversion of attention away from activities critical for safe driving toward a competing activity”. Technologies, such as hands-free voice commands, steering wheel controls, and mobile device integration allow for an increase in secondary tasks while driving, but these in-vehicle information systems (IVIS) impose variable amounts of cognitive and visual workloads on the driver. The visual and cognitive workloads induced by IVIS may have independent, additive, interactive, or multiplicative effects upon driver distraction.

The purview of the current study seeks to test novel measurements in the ability to detect components of overall workload. The current measurement devices and methods attempt to extend and refine reaction time (RT) measures currently being considered by the International Standards Organization (ISO) to standardize a measurement of RT protocol in vehicles (ISO, 2012). In this study, we employ two types of visual DRT stimuli and one vibro-tactile DRT to assess cognitive and visual workload (ISO, 2012). DRT stimuli were simultaneously mounted on the body and on the dashboard with both locations having two stimulus variants. In the first configuration (Experiment 1), participants responded to an ISO standard flashing red light mounted on the head and the dashboard. In the second configuration (Experiment 2), participants responded to an ISO standard vibro-tactile device on the body, and a modified light fading from red to green and back mounted on the dashboard. Both experiments utilized a cognitive task, a visual task, a combination of both, and a baseline for comparison. By simultaneously employing one DRT stimulus on the body and one mounted on the dash board of the vehicle, we expected that the DRT stimulus on the body would be sensitive to cognitive demands independent of visual attention, while the dash mounted DRT stimulus would be sensitive to shifts in visual attention. We hypothesized slower reaction times (RT) and lower hit rates under visual loads to the dash mounted DRT in comparison to the body mounted DRT. We also expected an additive effect of cognitive and visual workloads, leading to slower RTs and lower hit rates while performing both types of tasks compared to just one or the other.

METHOD

Participants

After an Institutional Review Board approval, participants were recruited through flyers posted on campus bulletins or students enrolled in a psychology course at the University of Utah. Thirty-six participants 18 to 54 years old (*M* = 23.4) completed the study. All participants reported normal or corrected-to-normal visual acuity and normal color vision (Ishihara, 1993). Participants received $20 compensation upon completion of the one-hour study.

Materials

A 106 cm diagonal Samsung Television monitor (1920 x 1080 pixels) displayed the background task approximately 91 cm from the participant. An L3 communications MPRI Ship Analytics, Simulation

Technology Solutions Simulator, manufactured by I-Sim Corporation was used. However, the simulator only provided the background screen, dashboard, steering wheel, and seat and the screens did not display a realistic driving simulation. A 24 x 19 cm (2048 x 1536 pixels) Apple® iPad® Mini 3 was fixed approximately 28 degrees down and 28.7 degrees to the right of the background target location. The target location on the background task followed a path that had a normal distribution for the frequency of position, moving laterally across the screen. A Seeing Machines Fovio™ Eye Tracker was used to track participants' eyes. A rotary encoder attached to the steering wheel of the driving simulator recorded deviation from target in the background task. The iPad® distractor task stimuli were presented by custom software projected to the iPad® using the Duet application. Participants wore one of two reaction time assessment devices. The study employed either a vibro-tactile (VT) or head mounted (HM) visual device. Both devices presented a stimulus on the bodies of the participants and a stimulus placed upon the dash of the driving stimulator. These Detection Response Task (DRT) devices followed the specifications outlined in ISO WD 17488 rev 10.1 (ISO, 2012). For the HM device, an LED light was mounted to a stalk connected to a headband.

Tasks. Both experiments included the three following tasks: The tracking task, the search task, and the Operation Span Task (OSPAN).

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 iPad® search task participants were asked to respond by touching the largest green ball on a blue background with the index finger of their right hand. The ball chooser game contained a target with 10 distractors. The distractors each covered 5% of the iPad® screen while the target covered 7% of the screen. If the participant chose an incorrect ball a chime sound played. If the participant chose the correct ball a new set of distracters and a target were immediately displayed.

Participants completed an auditory version of the OSPAN task developed by Watson and Strayer (2010). In this task participants attempted to solve mathematical problems while remembering single syllable words in serial order. In the auditory OSPAN task, participants were asked to remember a series of two to five words that were interspersed with math-verification problems (e.g., given "[6 / 2] – 4 = 5?" – "ice" – "[4 x 1] + 4 = 5?" – "tree" – RECALL, the participant answers "false" and "true" to the math problems when they are presented and recall "cat" and "box" in the order in which they are presented when given the recall probe).

Procedure

Participants in Experiment 1 were assigned to use the ISO standard head mounted and remote mounted visual DRTs for the duration of the experiment. In Experiment 2, participants were assigned to use the vibro-tactile body mounted and modified fading remote DRTs. Both experiments were conducted across four counterbalanced blocks for 8 minutes which all included the tracking task and responding to the DRTs. These tasks alone were considered the single-task block (ST). Participants also completed three dual task blocks, which included either the Operation Span (OSPAN) task to induce a cognitive load (C), a modified Triesman Conjunction Search task to induce a visual load (V), or both (CV) (Treisman & Gelade, 1980). 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 received a break between blocks.

In Experiment 1, the DRT tasks consisted of either a head mounted DRT (HDRT) or remote mounted DRT (RDRT). Participants responded to an LED light that flashed red for one second, mounted on either the dashboard, approximately 61 cm from the participants, or on a stalk attached by a headband up and to the left of the participant's visual field. The DRTs randomly presented a light every 3-5 seconds at either location. The stimuli were presented at one location at a time and the lights remained illuminated until a response was made or one second had elapsed. Response reaction time was recorded with millisecond accuracy. Participants responded by pressing a button attached to their left hand against the steering wheel.

In Experiment 2, participants responded to either an LED light mounted on the dashboard in the same location as Experiment 1 or an ISO standard vibro-tactile device (TDRT) taped just above the left collarbone. The remote mounted fading DRT (RFDRT) maintained a red light, which faded into green for one second or until the participant responded, and then faded back to red. Participants responded to either change by pressing the button attached to their left hand against the steering wheel. The light and the vibration remained engaged until a response was made or one second had elapsed. The Experiment 2 trials also occurred at one location every three to five seconds.

Measures. RTs to the presentation of a light or vibration stimulus 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. The Fovio™ Eye Tracker was calibrated to determine whether participants' eyes were directed to the forward screen or away from the forward screen. This was measured as a percentage of time eyes were on the forward screen. Steering error was calculated through deviation of the triangle from the ball using the rotary encoder. The measurements of eyes on the forward screen and steering error were used to validate DRT measures of visual load.

RESULTS

Experiment 1 – Headmounted Remote Flash DRT

Reaction Time. A repeated measures ANOVA determined a statistically significant difference in RT across conditions between head mounted (M = 642, SD = 32) and remote DRT (M = 702, SD = 28), $F(2, 25) = 51.146$, $p < .05$, $\eta^2 = .672$, demonstrating that remote DRT produced slower reaction times. There was also a significant difference between the four conditions, $F(2, 23) = 184.390$, $p < .05$, $\eta^2 = .96$. Finally a significant interaction between RT and Condition was found, $F(2, 23) = 3.133$, $p < .05$, $\eta^2 = .290$ (See Figure 1).

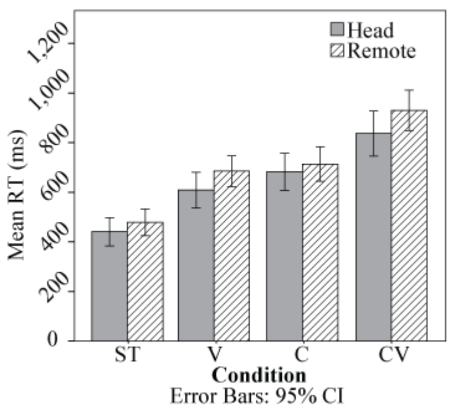

Figure 1. Response time in milliseconds to the head mounted and remote dash mounted lights across driving alone, visual, cognitive, and cognitive and visual conditions.

Hit Rate. A repeated measures ANOVA determined a statistically significant difference in hit rate between head mounted (M = 86%, SD = 4%, CI = (79%, 94%), n = 26) and remote DRT (M = 76%, SD = 4%, CI = (69%, 84%), n = 26), $F(2, 25) = 49.094$, $p < .01$, $\eta^2 = .663$, demonstrating that the RDRT resulted in fewer hits than the HDRT. There was also a significant difference between Conditions, $F(2, 23) = 10.787$, $p < .01$, $\eta^2 = .585$. Finally there was a significant interaction between Accuracy and Condition, $F(2, 23) = 3.271$, $p < .05$, $\eta^2 = .299$ (see Figure 2).

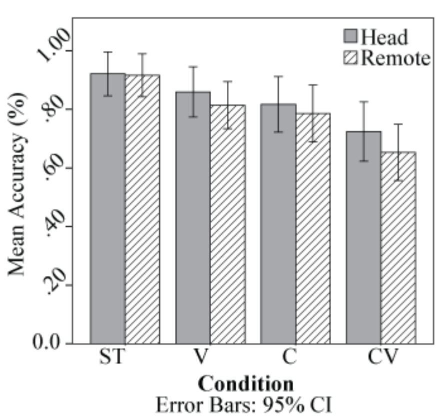

Figure 2. Percent accuracy of detection of the light stimulus in the head mounted and remote dash mounted lights across driving alone, visual, cognitive, and cognitive and visual conditions.

Steering Deviation. A repeated measures ANOVA ascertained the deviation from optimal tracking significantly differed across the four conditions $F(3,23) = 42.37$, $p < .05$ $\eta^2 = .847$ (See Figure 4). By comparing visual tasks (V, CV) and non-visual tasks (ST, C) in a post hoc analysis, participants performed better in non-visual tasks (M = 4.65, SD = 1.09) than in visual tasks (M = 7.34, SD = 1.76), $t(48) = -13.8$, $p < .05$, 95% CI[-2.29, -3.07] (see Figure 3).

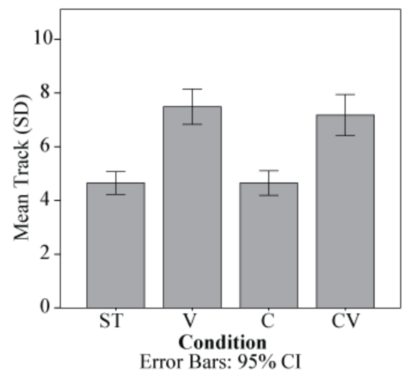

Figure 3. Deviation from tracking the optimal path of a ball across driving alone, visual, cognitive, and cognitive and visual conditions.

Eyes on Forward Roadway. A repeated measures ANOVA found that participants' eyes to the forward roadway differed across the four conditions $F(3,23) = 27.4$, $p < .05$ $\eta^2 = .805$. By comparing visual component tasks (V, CV) and non-visual component tasks (ST, C) in a post hoc analysis, participants' eyes in non-visual tasks (M = 97.9%, SD = 2.7%) were on the forward roadway more than in visual tasks (M = 74.7%, SD = 15%), $t(51) = -12.0$, $p < .05$, 95% CI[-19%, -26.6%] (see Figure 4).

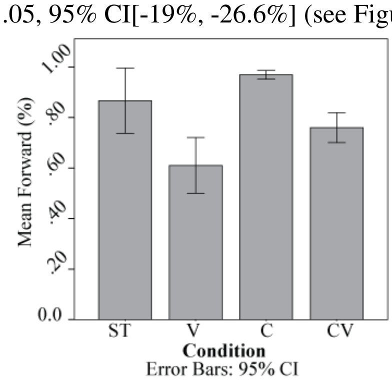

Figure 4. Percent time eyes are directed at the forward screen across driving alone, visual, cognitive, and cognitive and visual conditions.

Summary. Experiment 1 tested the sensitivity of a head mounted and remote mounted light configuration of DRT stimuli to cognitive and visual workload. Both ISO standard lights were predicted to be salient cues for participant responses in the simulator environment. Despite our prediction that simultaneously presenting two DRT stimuli in a novel configuration would be differently sensitive to cognitive and visual workload, both lights appeared to detect levels of visual and cognitive load in a similar manner. These data led to the conclusion that the Experiment 1 DRT configuration was unable to differentiate between cognitive and visual load. Due to the similarity of DRT stimuli in Experiment 1, Experiment 2 employed a vibro-tactile and remote mounted fading light to assess cognitive and visual workload. The TDRT was predicted to drive bottom up attention to the stimulus more strongly than a light regardless of visual load while it maintained sensitivity to cognitive load. The RFDRT was designed to be less salient than the ISO standard RDRT, and therefore more sensitive to changes in visual load. We predicted that both changes would lead to a larger difference in RTs and hit rates to the two types of DRTs, thereby unveiling more accurate separate measurements of visual and cognitive load.

Experiment 2 – Vibro-tactile Remote Fade DRT

Reaction Time. As in Study 1a, repeated measures ANOVAs displayed that responses to the vibro-tactile DRT (M = 553.6, SD = 31.3) were significantly faster than to the remote mounted fading light (RFDRT) (M = 842.1, SD = 25.9), $F(1, 23) = 68.115$, $p < .001$, $\eta^2 = .748$. Additionally, RTs in each condition differed significantly across condition, $F(2, 21) = 48.03$, $p < .001$, $\eta^2 = .873$. However, an interaction between the location of the stimulus and the condition was not statistically significant (see Figure 5).

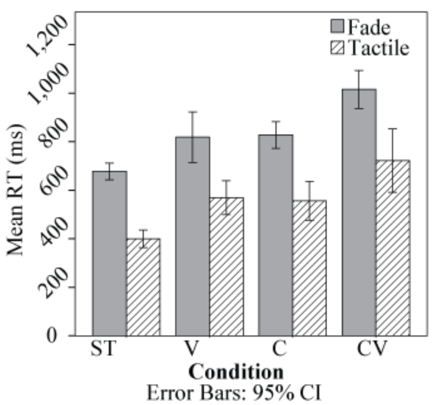

Figure 5. Response time in milliseconds to the vibro-tactile and remote mounted fading light across driving alone, visual, cognitive, and cognitive and visual conditions.

Hit Rate. A repeated measures ANOVA revealed that hit rate significantly differed between the VT (M = 91.2%, SD = 11.5%) and the RF light (M = 70.7%, SD = 23.7%) across conditions, $F(1, 24) = 31.8$, $p < .001$, $\eta^2 = .57$. Accuracy also differed significantly across the four conditions, $F(3,22) = 20.7$, $p < .001$, $\eta^2 = .74$. Finally, there was a location and device type by condition interaction for accuracy, $F(3,22) = 28.88$, $p < .001$, $\eta^2 = .797$ (see Figure 6). This interaction suggests a strong effect of DRT type across the different conditions, namely that hit rates decreased to the RFDRT in visual conditions while they did not for the TDRT.

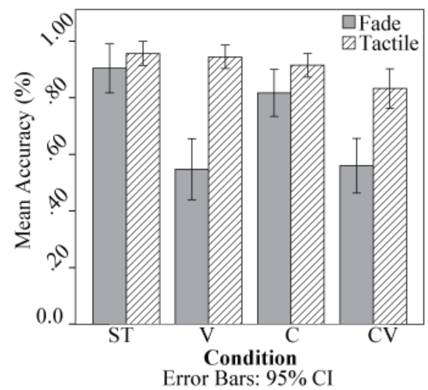

Figure 6. Percent accuracy of detection of the stimulus in the vibro-tactile DRT and the remote mounted fading DRT across driving alone, visual, cognitive, and cognitive and visual conditions.

Steering Deviation. In a repeated measures ANOVA, optimal tracking in the tracking task significantly differed across the four conditions, $F(3,21) = 32.76$, $p < .05$ $\eta^2 = .824$. By comparing tasks that have a visual component (V, CV) and tasks that do not have a visual component (ST, C) in a post hoc analysis, participants in non-visual tasks (M = 3.87, SD = 2.16) are significantly better at tracking than in visual tasks (M = 6.96, SD = 1.18), $t(47) = 12.2$, $p < .05$, 95% CI[2.59, 3.6] (see Figure 7).

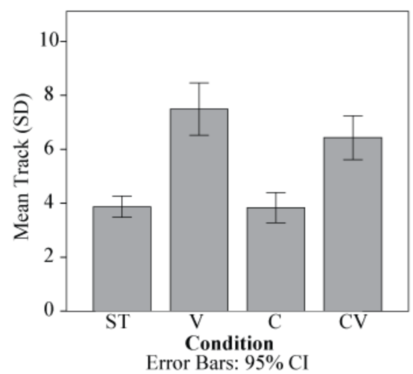

Figure 7. Deviation from tracking the optimal path of a ball across driving alone, visual, cognitive, and cognitive and visual conditions.

Eyes on Forward Roadway. A repeated measures ANOVA revealed that participants spent differing amounts of time with their eyes on the forward roadway across the four conditions $F(3,20) = 28.01$, $p < .05$ $\eta^2 = 0.8$. In visual tasks (V, CV) compared to non-visual tasks (ST, C) in a post hoc analysis, participants' eyes in non-visual tasks (M = 91%, SD = 20%) were forward more than in visual tasks (M = 72%, SD = 21%), $t(46) = -8.34$, $p < .05$, 95% CI[-23.8%, -14.6%] (see Figure 8).

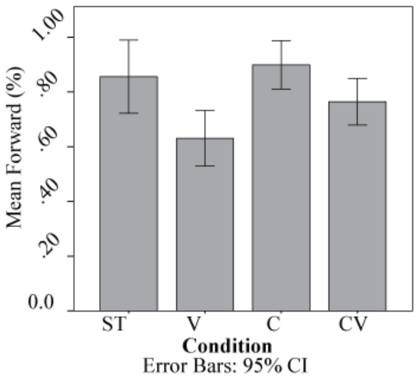

Figure 8. Percent time eyes are directed at the forward screen across driving alone, visual, cognitive, and cognitive and visual conditions.

DISCUSSION

This study attempted to develop and refine the efficacy of RT and hit rate measures in four types of DRTs to simultaneously distinguish between levels of visual and cognitive workload. In order to validate whether the DRTs were independently and simultaneously sensitive to eye glances and cognitive load, we compared RTs and hit rates across four conditions.

In Experiment 1, both DRT devices seemed to be sensitive to the relative level of mental workload across the four conditions, but did not seem to differentiate RT or hit rate in response to visual and cognitive load separately. This would suggest that RTs and hits to the light that stayed within the visual field and the light mounted on the dashboard of the simulator performed similarly in detecting visual and cognitive workload.

In Experiment 2, the remote fading DRT (RFDRT) seemed to be sensitive to visual workload in hit rate relative to the TDRT, but not in RT. This finding suggests RT and hit rate may be sensitive to different processes in measuring visual and cognitive workload. The patterns of mean deviation in the tracking task and eye-tracking data for time spent looking to the forward roadway further support the RFDRT's sensitivity to eye location in Experiment 2. As expected, the percentage of time eyes were fixed on the forward roadway was significantly less during visual tasks than during non-visual tasks, which mirrors the pattern of hits for the RFDRT (see Figure 6 & Figure 8 for comparisons). The eye-tracking measure of percent of time eyes were fixed on the forward roadway validates the significantly different hit rates for the RFDRT between visual and non-visual tasks. Therefore, the hit rate to the RFDRT seems to be a sensitive surrogate for visual attention during a driving task, and due to its simplicity can be readily used as a measure of visual load in the vehicle.

For the TDRT, the vibro-tactile device seems to be able to cut through distractors in visual attention and remain relatively salient to participants. Despite finding significant differences across conditions in RT and hit rate to the TDRT, the TDRT appears to be less sensitive to different levels of visual demand in hit rate than the RFDRT. Because RTs and hit rates to the vibro-tactile device did not seem to be affected by the visual tasks in comparison to the RFDRT, we may conclude that the changes across conditions are a measure of cognitive load.

Overall, this study attempted to test DRT configurations that would simultaneously assess cognitive and visual workload during a simulated driving task. Experiment 2 provides evidence that assessing these components of workload simultaneously in the vehicle is feasible with this modified approach. The combination of a vibro-tactile device and the modified fading light provide promising response measures that may be used to update the current ISO standards of driver workload assessment in the vehicle, and further distinguish between different components of overall workload.

REFERENCES

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. (2012). 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.

Ishihara, S. (1993). Ishihara's tests for colour-blindness. Tokyo, Japan: Kanehara & Co.

NHTSA. (2012). Visual-Manual NHTSA Driver Distraction Guidelines for In-Vehcile Electronic Devices. Department of Transportation. Docket No. NHTSA-2010-0053. Paris, C. R., Gilson, R. D., Thomas, M. H., & Silver, N. C. (1995). Effect of synthetic voice intelligibility on speech comprehension. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37(2), 335-340.*

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.

Watson, J. M., & Strayer, D. L. (2010). Supertaskers: Profiles in extraordinary multitasking ability. Psychonomic bulletin & review, 17(4), 479-485.

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

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta psychologica, 41*(1), 67-87.