The Driving Simulator as a Workload Research Tool

Justin F. Morgan, Tal Oron-Gilad, Peter A. Hancock

Department of Psychology, PO Box 161390, Orlando, FL 32816-1390

Abstract

This paper shows that the driving simulator may be used as a platform in which to study workload and workload performance relationship. The participants were 8 adults with valid driver's licenses. After a familiarization drive participants drove a route with five equidistant and symmetrical segments using a map and turn-by-turn directions. Upon completion participants retraced their journey without the use of the map or directions. Measures of workload and simulation sickness were collected pre-immersion and after completion of each of the two driving phases. Significant difference in workload between conditions were present. No significant results from the simulator sickness questionnaire (SSQ) were present. These findings suggest that the task was the primary contributor to workload, and support the use of driving and driving simulators as a workload research platform.

Introduction

The task of driving is inherently stressful. There are large demands made of the driver's cognitive abilities, as well as physiological demands that controlling the vehicle require. Driving is generally considered a high workload task, as evidenced by the large amount of research performed examining the mental workload aspects of driving (for a review see Recarte & Nunes, 2003).

Increasing mental workload on the part of the driver is associated with longer reaction times to signals such as changing traffic lights and oncoming traffic (Patten, Kircher, Östlund, & Nilsson, 2004). In fact, many types of in-vehicle displays have been shown to lower overall driver performance (Lansdown, Brook-Carter, & Kersloot, 2004). In general, increasing drivers' mental workload has lead to a reduction in objective measures of performance.

Although there is often disagreement about the definition and boundaries of the topic, workload is one of the most commonly studied topics in human performance. Over the decades, many assessment techniques have been developed for measuring human workload. Common objective and subjective measures of workload take into account factors such as mental workload, temporal demands, and effort in a particular task to generate an overall description of workload. These measures have proven to represent the workload experienced by the operator in a variety of settings and conditions.

Although workload research has occurred in a variety of settings, little research has been performed which uses the concept of driving, and in particular driving simulation, as a platform for workload research. The present study seeks to provide data which supports the use of a driving simulator for workload research by demonstrating that reliable and accurate measures from a well-regarded workload assessment instrument are available in a simulated driving environment and that changes in workload observed during the experiment were not due to physiological changes due to discomfort from the simulator.

Method

Participants

Eight adults (four males and four female) participated in this experiment. The five participants had an average age of 24.4 years (*SD* = 4.4). No particular inclusion/exclusion criteria were set for participation besides holding a valid driver's license, having normal color vision, and normal or corrected-to-normal vision. Participants were compensated for their participation with a payment voucher which can be exchanged for $7.50.

Simulator

One PatrolSim Mark III driving simulator was used for this experiment. The PatrolSim is a fixed-base, medium fidelity, driving simulator manufactured by L3. The simulator has all standard cabin controls of an automobile (steering, brake, throttle, turn signals, etc.), and sits in front of a projection screen which yields an area of approximately 150° horizontally which the participants were seated approximately three feet from.

Workload Measure

The NASA Task Load Index, or TLX, (Hart & Staveland, 1988) was administered via computer to each participant. The TLX is a multidimensional scale which measures six aspects of perceived workload (mental demand, physical demand, temporal demand, performance, effort, and frustration) on a scale from 0 to 100. In addition to the six individual scores and their weights, an overall value of perceived workload is generated. For the purpose of this research, only the combined, total workload score is examined.

Simulation Sickness Measure

Developed by Kennedy, et al. (1993), the Simulator Sickness Questionnaire (SSQ) is a widely used multidimensional measure of simulation sickness. The SSQ uses a pre-immersion interview and symptom checklist as a baseline to which subsequent symptom checklists are compared to. The SSQ measures symptoms clustered into three dimensions: nausea, oculomotor, and disorientation. For the purposes of this analysis, only the combined, total score is examined.

Procedure

After obtaining informed consent, the participant was administered the SSQ pre-immersion interview and symptom checklist forms. Following completion of these, the participant completed a baseline measure for the NASA TLX. Participants were then asked to enter the simulator. Each participant was first given a five minute period to familiarize them with the operation of the simulator. This initial exposure to the simulator was done in order to reduce the likelihood of the participant experiencing symptoms of simulator sickness and to reduce the participant's uncertainty with the control of the simulator.

Participants were then given a map and a set of turn-by-turn driving directions. Participants were instructed to proceed according to the directions to a series of six markers placed at approximately equal distances in the environment. When they reached each marker, participants were instructed to come to a complete stop, honk the automobile's horn, and then continue to the next marker. Participants were informed that they could refer to the map and driving directions at any point in the task. After completing the driving route, post task versions of the SSQ and NASA TLX were administered.

For the second driving task, participants began from the last marker and were asked to navigate through the markers in reverse order, without the aid of a map or driving directions. Participants were instructed to continue as far as possible and that when they felt they were lost or unable to find the next marker, inform the experimenter of this and the experiment would stop. At completion of the route or after the participant could not find the next point, post task versions of the SSQ and NASA TLX were administered and the participant was debriefed.

Results

Completion of the task

All participants completed the entire route in the map condition. The number of participants who completed the no map condition declined as the task continued.

	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
Participants Remaining	8	7	6	4	2
Percent of Participants Remaining	100	87.5	75	50	25

Table 1: Participant completion of the no map condition

When examined graphically, the percentage of participants completing the no map condition demonstrated a sharp decline (see figure 1).

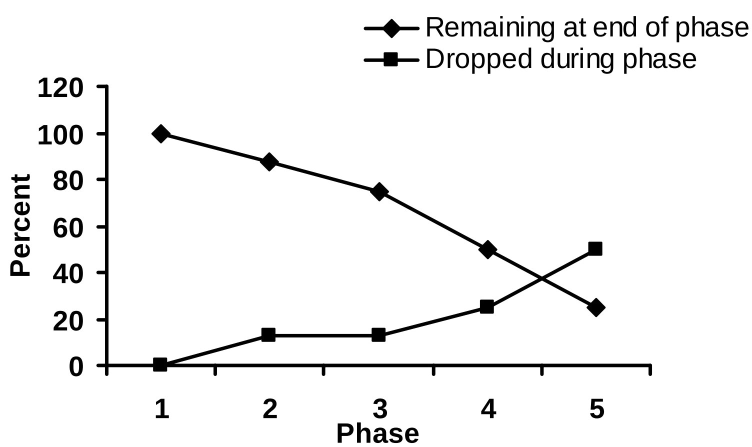

Figure 1: Percentage of participants dropped and remaining during no-map phase.

NASA TLX

A significant change was present in the workload scores across the three phases (practice session, driving with map, driving without map), $F(2,21) = 11.358$, $p < .0005$, part. $\eta^2 = .520$. Planned comparison analysis shows that the map ($M = 55.83$, $SD = 13.79$) and no map ($M = 72.00$, $SD = 4.49$) phases differed significantly from one another, while the baseline ($M = 38.50$, $SD = 19.56$) differed significantly only from the no map phase.

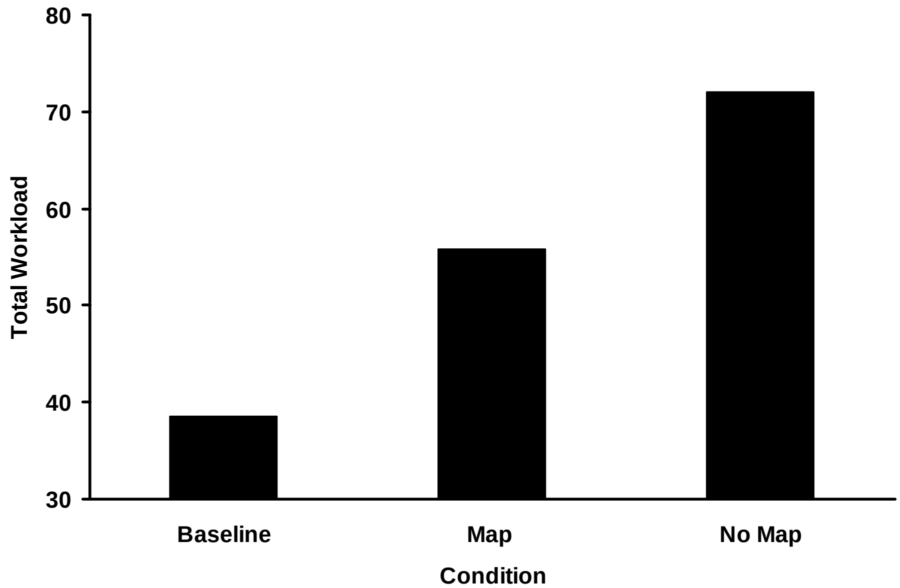

Figure 2: NASA TLX Scores by phase of participant drive

SSQ

There were no significant effects were present for the data from the SSQ, $F(2,21) = 2.264$, $p = .129$, part. $\eta^2 = .177$, an increase in SSQ total scores was observed between the baseline measure (M = 57.49, SD = 84.89), the first post-immersion measure after the map condition (M = 179.11, SD = 133.24) and the last measurement after the no map condition (M = 199.21, SD = 193.29).

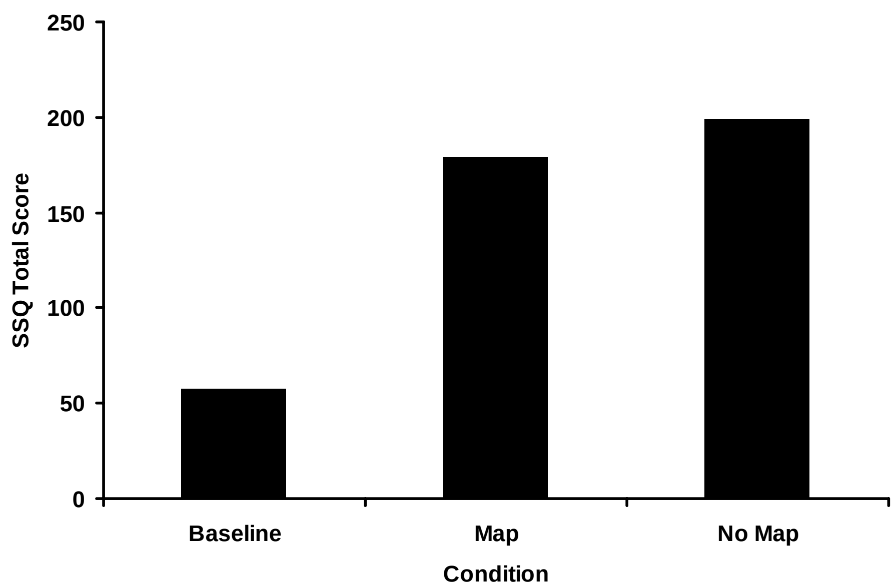

Figure 3: SSQ Scores by phase of participant drive

Discussion

The lack of any significant differences between time measures for driving different segments of the route demonstrated that the route was symmetrical and equidistant in layout. This allowed for the examination of the workload measures as a measure of the navigational task instead of a measure of the driving task. The lack of a significant difference between baseline and map driving conditions on the workload measure show that the actual task of driving, when no additional stressors are introduced, is not a significant contributor to workload. The addition of an additional workload item, in this case the navigation without a map, greatly increases workload over both baseline and map conditions. These support the hypothesis that the navigational task was the significant contributor to workload in this experiment.

Participants, on average, did not experience many symptoms of simulator sickness during the experiment. This is supported by the gradual, yet non-statistically significant, increase in scores on the SSQ. The SSQ displays a wide range of possible scores, and is extremely sensitive to changes in participant symptoms. This led to wide variability in scores, which likely contributed to the non-significance of the differences in conditions. However, the lack of a significant increase in the SSQ leads to the conclusion that workload was not being reliably influenced by any symptoms of simulator sickness, and was instead influenced solely by the navigational task.

Of interest in this research project is the validation of a ready to use driving simulator for workload research. We attempted to demonstrate this by showing that while driving an equidistant and symmetrical route under varying degrees of workload, high quality measures, which were due to the task characteristics, could be obtained and that the task's workload could be reliably separated from the workload imposed by driving. This was supported by the experimental findings, and supports the idea that a driving simulator can be used as a platform for workload research.

References

Hart, S.G. & Staveland, L.E. (1988). Development of the NASA TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North-Holland.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. International Journal of Aviation Psychology, 3, 203-220.

Lansdown, T.C., Brook-Carter, N., & Kersloot, T. (2004). Distraction from multiple in-vehicle secondary tasks: vehicle performance and mental workload implications. Ergonomics, 47, 91-104.

Patten, C.J.D., Kircher, A., Östlund, J., & Nilsson, L. (2004). Using mobile telephones: Cognitive workload and attention resource allocation. Accident Analysis & Prevention, 36, 341-350.

Recarte, M.A., & Nunes, L.M. (2003). Mental workload while driving: Effects on visual search, discrimination, and decision making. Journal of Experimental Psychology: Applied, 9, 119-137.