RP-95-48

CHALLENGES TO THE DESIGN OF DRIVING SIMULATOR EXPERIMENTS

Professor Paul P. JOVANIS

Associate Director
Institute of Transportation Studies

Professor
Civil and Environmental Engineering
UNIVERSITY of CALIFORNIA at DAVIS

Résumé

Il ne fait plus de doute que la simulation de conduite est un outil de plus en plus accessible et utilisée pour l'étude des problèmes de transport. Des simulateurs de conduite à la pointe sont maintenant disponibles en Europe, aux Etats-Unis et au Japon.

Pourtant, la simulation de conduite est remise en question par ceux qui arguent son coût excessif et ses résultats limités.

Un des principaux défis de la simulation réside dans la définition des variables à mesurer et analyser.

Ce papier a pour principal but d'argumenter le fait que des expériences nouvelles et inventives doivent être menées pour utiliser des méthodes d'analyse statistique avancées.

Abstract

In every study referenced in this paper, ANOVA or regression was used to study a single dependent variable. Several studies used more than one measure, but the analysis of each was separate.

It is the author's contention that this comes, in part, from the strong experimental training that is provided to most researchers in the field. Control of experiments is emphasized. Rigi independent variables are developed and used in the experimental design with a simple dependen measure. Simulator studies will be much more effective, useful, and powerful if multipl measures are linked.

Two examples of linked dependent measures are given in the paper. The dependency of spee and headway, (or speed and reaction time), eye movement, and lateral position. Statistic procedures such as variable instrumentation have been used in survey data (e.g. Abdel-Aty, et and are directly applicable here. What is more difficult is to design experiments to take advant of these methods. This requires training of the experimenters and some courage in conduct more flexible experiments.

l de plus en plus
s simulateurs de
-Unis et au Japon.

i arguent son coût

n des variables à

nces nouvelles et
alyse statistique

is used to study a single
the analysis of each was

xperimental training that
s is emphasized. Rigid
with a simple dependent
d powerful if multiple

The dependency of speed
ral position. Statistical
a (e.g. Abdel-Aty, et. al.)
iments to take advantage
e courage in conducting

There is no question that driving simulation is an increasingly available and utilized tool for the study of a variety of transportation problems. Very advanced driving simulators are now available at several sites across Europe (e.g. Daimler Beuz and VTI), the United States (e.g. University of Iowa, U.S. Federal Highway Administration, General Motors, Hughes Aircraft Co.) and Japan (several auto manufacturers). In addition, increases in computing power make available and affordable a variety of part-task desk top simulators that can also be used for a variety of studies (e.g. Srinivasan, et. al., 1995).

Despite this proliferation, driving simulation faces serious challenges from those who believe that the tool is too expensive and the results are too limited (some may say simplistic). Significant time must be devoted to developing driving scenes, scenarios and variable measurements, customized for almost every experiment. Many have challenged the need for advanced simulators, arguing that much relevant knowledge can be learned in simpler, cheaper experiments.

One of the difficult challenges posed by driving simulation is the question of which variables to measure and analyze. Most simulator studies are concerned with how well the driver undertakes the driving task when confronted with a set of devices (e.g. navigation displays) or a challenging driving environment. Measures of safety or accident risk are frequently the underlying objective. Real accidents, however, do not usually result from a single factor but multiple factors. It can therefore be argued that multiple measures can and should be used to assess safety in a simulator environment. Clearly, the environment supports and encourages such data collection; a host of measures are available in a simulator much more readily than is on the road.

The principal argument of this paper is that new and more imaginative experiments must be conducted using more advanced statistical analysis methods. The statistical techniques should allow for multiple dependent measures considered jointly not separately. This is the principal value added from advanced driving simulation: more valid experiments with multiple performance measures in a complex driving environments.

The paper proceeds by reviewing the literature concerning simulator studies, identifying measures used, and statistical procedures employed. The vast majority of studies use a single measure and utilize ANOVA to identify effects of a set of independent variables on the dependent variable. The review is organized by the dependent variable used in each study. Periodically examples of richer multivariate analyses are identified.

Dependent Variables Used in Simulator Studies

Driving performance is measured by one or more of several dependent variables. The measures of driving performance that have been typically used in human behavioral studies include: car following headway, lateral position, accuracy in using the information system, time of reaction and execution thereafter, number of navigation errors, eye movement, glance duration, visual performance, heart rate, vehicle speed and/or acceleration, and accidents.

Car Following Headway

Headway, defined as the distance from the leading car to the test car, is one way to measure the effectiveness and safety of driving and is commonly used in evaluating driver performance. It has been measured in terms of headway (Noy, 1990), heading error (Ranney and Gawron, 1986a and 1986b), and heading error rate (Weir and Wojcik, 1971). The headway is computed from the speed and acceleration curves (ultimately, the trajectories) of the lead vehicles and those following. Often, subjects are asked to maintain a certain headway while following a vehicle. Maintenance of a smaller headway could mean that the subject is not following at a safe distance (risking an accident), while a larger headway may be an inefficient use of road space. Considering headway alone, however, may not be very useful; the speed of the vehicle and the reaction time of the subjects must also be considered. For example, a smaller headway may be maintained by a driver if their reaction time is short of if traveling at a slower speed. The studies by Noy, and Ranney and Gawron, used ANOVA to analyze each dependent measure separately. One improvement is to include predicted speed as an instrumented variable in the headway equation. As speed and headway are endogenous (i.e. chosen by driver) this is a proper formulation and could control for dependencies.

Lateral Position

Lateral position is defined as the position of the imaginative center line of the car or the track of the wheels relative to the center line of the lane. Lateral position is used as a safety indicator, as any substantial deviation from the lane center could lead to the possibility of a collision with cars in an adjacent lane. Lane keeping has been evaluated in terms of road or lateral position error (Sussman et al., 1971); Ranney and Gawron, 1986a, 1986b), the amount of road used or the lateral position itself, lane boundary exceedences (Dingus et al., 1989), lateral acceleration (Ranney and Gawron, 1986a), lateral velocity, lateral deviation and heading angle (Weir and Wojcik, 1971), lane departure frequency (Gawron and Ranney, 1988; Imbeau et al., 1989; Noy, 1990), lane exceedence ratio (Noy, 1990), average and variance of lateral placement (Walker et al., 1990). ANOVA was commonly used in these analyses as well.

Accuracy (Error Rate or Error Frequency)

This is measured in terms of the information missed or interpreted wrongly as well as the resultant errors committed by a driver in completing a certain task. In some of the studies, subjects are given points for each of these errors and the subject with the greatest number of points is designated as being at greater risk than the other drivers. Different studies have used different criteria to measure this performance. Number of task errors (Dingus et al., 1989), number of obstacles struck (Ranney and Gawron, 1986a 1986b), sub of the tracking errors (Public Service Research Inc., 1961), frequency of navigational errors (Walker et al., 1990) and percentage of correct recognition of message on display (Emmerson and Linfield, 1986), are some examples. The subject should be able to obtain the information provided and be able to complete the task correctly and timely (Dingus et al., 1989). This measures the efficiency of the information system in terms of how long it takes the driver to react to the cues from the in-vehicle navigation systems, to find the correct intersections (upcoming approaches) to a destination, as well as the total time required to complete the trip.

Reaction Time

The time elapsed before the driver reacts to a stimulus is measured as the reaction time or the response time. The stimulus can be in the form of traffic markings, signs and signals which assist in the driving task, a secondary task, or a movement of the lead vehicle (slowing down or stopping). This driving performance measure is particularly important when the subject has a large reaction time, as one may not be able to respond in time to prevent an accident. A short response time usually translates into higher efficiency and a higher degree of safety (more accurate responses leading to prevention of potential accidents).

Ranney and Gawron (1986a) used the standard deviation of reaction time to a sign, along with the mean detection time and studied their variation across individual driver's different situation demands. Noy (1990) used reaction time to the auxiliary task as a performance measure. Walker et al., (1990) used reaction time to gauge changes as one of the performance measures in the study.

Eye Movement

The visual demands inside and outside the vehicle influence eye movements and glance duration, which are good performance indices of the effectiveness and safety of driving (Imbeau

way may be maintained by a driver
The studies by Noy, and Ranney
separately. One improvement is to
ieadway equation. As speed and
r formulation and could control for

center line of the car or the track of
ion is used as a safety indicator, as
possibility of a collision with cars in
is of road or lateral position error
), the amount of road used or the
, 1989), lateral acceleration (Ranney
heading angle (Weir and Wojcik,
; Imbeau et al., 1989; Noy, 1990),
eral placement (Walker et al., 1990),

interpreted wrongly as well as the
task. In some of the studies, subjects
ith the greatest number of points is
Different studies have used different
rs (Dingus et al., 1989), number of
f the tracking errors (Public Service
lker et al., 1990) and percentage of
Linfield, 1986), are some examples
ed and be able to complete the task
efficiency of the information system
om the in-vehicle navigation system
destination, as well as the total time

measured as the reaction time or the
rkings, signs and signals which assist
the lead vehicle (slowing down
important when the subject has a large
revent an accident. A short response
ree of safety (more accurate response

of reaction time to a sign, along with
dividual driver's different situations
sk as a performance measure. Walker
of the performance measures in the

nfluence eye movements and gl
veness and safety of driving (Imbe

---

al., 1989). Less complex driving surroundings, secondary tasks or information system will allow drivers to pay more attention to following the lead car. Hence, they can drive in a safer way, maintaining a reasonable headway and staying within the lane. The information presented on the display should thus be simple and easy to understand and should be presented for a reasonable duration, so that the driver needs less time to respond and take any necessary action. Simplicity in the display, then, will have a beneficial effect on reaction time. Large character size and high illumination level also decrease response time. While eye movements and glance duration have been measured by Miura (1986) for actual driving situations, Imbeau et al., (1989) and Dingus et al., (1989) considered the gaze duration (time spent looking at the display) and its effects on the other driver performance measures. Noy (1990) used look frequency, dwell time and viewing ratio. Duration of glances to the center roadway or the navigation display and the frequency of transitions between the roadway center, mirrors and dashboard, have been used as performance measures in a field study of a particular display (Wierwille et al., 1987).

Vehicle Speed

Speed can be a confounding factor or a dependent variable in driver behavior studies. It may relate to the degree to which drivers are willing to take a risk (Matthews, 1986; Jonah and Dawson, 1987). Speed may also affect the relevant reaction time. But it may not be correlated with accidents (Harms, 1986). It is used to calculate the car's trajectory which in turn is used to measure the headway. While negotiating a curve, the speed and acceleration of the lead car will have an important impact on drivers' behavior, affecting other performance measures like reaction time, lateral position and headway (Ranney and Gawron, 1986a). Thus, it could be treated as an intermediate dependent variable in any study relating to other dependent variables.

Accidents

One of the principle objectives of driver behavior studies is to enhance safety be reducing the potential for accidents. Accidents can occur due to operating errors in any of the areas mentioned above (appropriate headway, suitable lateral position, accuracy of following instructions from on-board or outside information systems, and response time). While it is not possible to have a real accident in a simulation experiment, simulated accidents can be measured and do occur, particularly when the driving task has high demand. While one would expect subjects to generally avoid accidents, even in the simulator, their occurrence can be used as a direct measure of driving performance.

Mental Workload Measures

There has been no universally accepted definition of mental workload. According to Sheridan (1990), mental workload can be defined and measured in four different ways: (i) Secondary task performance; (ii) Measures of primary task complexity, independent of experimental subject; (iii) Physiological indices; and, (iv) Subjective workload measures.

Each of these measures four methods is discussed briefly below.

Secondary Task Performance

In this method, the subject is asked to perform a secondary task at the same time he is performing the primary task. This method is based on the argument that, the better performance on the secondary task, the less the workload of the primary task (Sheridan, 1990).

Walker et al., (1990) used this method in their simulator study. Subjects were asked to solve mental arithmetic problems while driving. They were also asked to monitor and respond to changes in two gauges by pushing buttons that were located in the dashboard.

Measures of Primary Task Complexity

This is a measure of a task, not the subject. Examples may be; number of steps to be undergone, improbability of the average step, etc. Apparently, this measure has been used for aviation studies (Sheridan, 1990).

Physiological Measures

Some of the physiological measures that have been used in previous studies are: heart rate, heart rate variability, pulse rate, electrical resistance changes of skin due to incipient sweating, pupil diameter and changes in breathing pattern (Sheridan, 1990). There are some disadvantages in using these measures. Physiological measures have been found to be sensitive to noise and emotion induced effects. They have also been found to be sensitive to large inter-individual differences (Verwey, 1990). No single measure has been accepted as a standard. However, there seems to be a preference for using heart rate/heart rate variability as an indicator for measuring general workload.

Hahn and Kading (1988) used speed, loss of control (times), reaction time, pulse rate, respiration rate and skin temperature as performance indicators in their study on 3 different types of steering controls (conventional rear wheel drive, permanent 4 wheel drive and 4MATIC -'self activating four wheel drive). 4MATIC drivers 'lost control' the least number of times, and had the lowest respiration and pulse rates. Similarly, permanent four wheel drivers lost control the most number of times, and had the highest respiration and pulse rates. Walker et al., (1990) used heart rate as one of the dependent variables in their study. However, the results indicated that heart rate was neither responsive to loading nor to the complexity of the navigational displays.

Subjective Measures

According to Sheridan (1990), subjective measures are the standard against which all the objective measures of workload are compared. Examples of some popular subjective workload measures are: Subjective Workload Assessment Test (SWAT), NASA Task Load Index (NASA TLX), and the Cooper Harper rating scales. Apparently, the SWAT and NASA TLX methods are the two methods that are being used most often.

Abdel-Aty, M
(1994). "Mo
Southern Cal
(1994)

Dingus, T.A.
Navigation Sy

Emmerson, F
Vision in Veh
57-65.

Hahn, Stefan
of Selected Ex

Harms, Lisbe
Task Real Tra
B.V. (North-H

Imbeau, Danie
Instrumental F
Human Factor

Miura, Toshia
Peripheral Vis
Publishers B.V

Noy, Ian Y (
Displays". Ro
No. TP 10727

Public Service
Task Situation
Department of

Ranney, T.A.,
a Driving Sim
525.

Ranney, T.A.
Driving Perfor
269.

be; number of steps to be
measure has been used for

ious studies are: heart rate,
n due to incipient sweating,
here are some disadvantages
to be sensitive to noise and
ve to large inter-individual
a standard. However, there
an indicator for measuring

, reaction time, pulse rate,
study on 3 different types of
l drive and 4MATIC - self
umber of times, and had the
lrivers lost control the most
ker et al., (1990) used heart
ults indicated that heart rate
onal displays.

ndard against which all the
opular subjective workload
A Task Load Index (NASA
nd NASA TLX methods are

REFERENCES

Abdel-Aty, Mohamed, Kenneth Vaughn, Ryuichi Kitamura, Paul P. Jovanis, and Fred Mannering (1994). "Models of Commuters' Information Use and Route Choice: Initial Results Based on Southern California Commuter Route Choice Survey". Transportation Research Record No. 1453 (1994)

Dingus, T.A., et al. (1989). "Attentional Demand Requirement of an Automobile Moving-Map Navigation System." Transportation Research A, 23A(4): 301-315.

Emmerson, P.G. and P.B. Linfield (1986). "Perception of Variable Message Traffic Signs." Vision in Vehicles, Edited by A.G. Gale et al., North Holland: Elsevier Science Publishers B.V., 57-65.

Hahn, Stefan and Wilfried, Kading (1988). "The Diamler-Benz Driving Simulator - Presentation of Selected Experiments". SAE Paper 880058, Warrendale, PA.

Harms, Lisbeth (1986). "Drivers' Attentional Responses to Environmental Variations: A Dual-Task Real Traffic Study." Vision in Vehicles, Edited by A.G. Gale, Elsevier Science Publishers B.V. (North-Holland), 131-138.

Imbeau, Daniel, Wierwille, Walter W., Laurie D. Wolf and Gail A. Chun (1989). "Effects of Instrumental Panel Luminance and Chromaticity on Reading Preference in Simulated Driving." Human Factors, 31(2): 147-160.

Miura, Toshiaki (1986). "Coping with Situational Demands: A Study of Eye Movements and Peripheral Vision Performance." Vision in Vehicles, Edited by A.G. Gale et al., Elsevier Science Publishers B.V. (North-Holland), 205-221.

Noy, Ian Y (1990). "Attention and Performance while Driving with Auxiliary In-Vehicle Displays". Road Safety and Motor Vehicle Regulation - Transport Canada, Ottawa. Publication No. TP 10727 (E), December.

Public Service Research Inc. (1961). A Study of Performance Decrement in Simulated Complex Task Situations and Simulated Environment. Prepared for Division of Accident Research, U.S. Department of Health, Education and Welfare, Washington D.C.

Ranney, T.A., and V.J. Gawron (1986a). "The Effects of Pavement Edgelines on Performance in a Driving Simulator under Sober and Alcohol-Dosed Conditions." Human Factors, 28(5): 511-525.

Ranney, T.A. and V.J. Gawron (1986b). "Task Demand and Alcohol Effects on Simulated Driving Performance." Proceedings of the Human Factors Society - 30th Annual meeting, 265-269.