# Driver performance evaluation: Considerations underlying selection and design of routes

M.Sodhi[1], B.Reimer[2], M.Tant[3], E.Peli[4],A.Bowers[4],R.Woods[4],K.Higgins[5] and P.Turco[6]

[1]University of Rhode Island, Kingston, RI, USA
[2]Transportation Center, Massachusetts Institute of Technology, Cambridge, USA
[3]Belgian Road Safety Institute (BIVV/CARA), Brussels, Belgium
[4]The Schepens Eye Research Institute, Harvard Medical School, Boston, MA, USA
[5]The Lighthouse International, New York,USA
[6]Massachusetts Eye and Ear Infirmary, Boston MA, USA

There is increasing interest in systematically studying the risks encountered while driving. In some cases, the focus is on potential risks such as those associated with the use of in-car devices (e.g., mobile phones, radio, navigational displays) or the effect of substances such as alcoholic beverages or medications. In such studies, the objective is in determining whether there is a decrement in driving performance that could lead to an increase in the probability of an accident, i.e. turns a safe driver into an unsafe driver. Another type of study involves the evaluation of the potential benefits/risks associated with the use of corrective/assistive devices (e.g. hand controls, vision enhancement devices) that might enable individuals with disabilities to drive. Here, the primary interest often is in determining if the device in question produces an increment in driving performance that could lead to a decrease in the probability of an accident, i.e., turns an unsafe driver into a safe driver. In both of the above cases, it is ultimately important to address the issue(s) through on-road testing. However, because of the different aims of the two types of studies, care must be taken to ensure that the design of the test route(s) and driving tasks are deliberately biased to maximize the likelihood of answering the critical questions. Two examples will be described to illustrate this point. The first will consider the special design elements to evaluate the impact of in-vehicle mobile phone use by normally-sighted drivers on overall driving performance. The second will consider the special design elements of a road test course to evaluate the effectiveness of vision devices as driving aids for visually-impaired people.

## 1. INTRODUCTION

Advances in vehicular technology have made driving safer now than ever before. However, this increase in safety has been offset by increasingly ageing driving populations, longer daily travel times and widespread use of mobile computing and communication devices while driving. Although in-vehicle devices such as the radio, climate controls etc. have been located on automobile dashboards since the first prototypes, it is possible that these distract the driver. In spite of the lack of strong supporting evidence, there is not much concern about the risk incurred by the routine use of these devices while driving. The current controversy and concern is centered on the risk posed by the use of recently introduced in-car devices such as cell phones, navigation systems etc. Another impact of the advances in technology relates to the level of assistance that can be offered to drivers with impairments. Vehicles customized for disabled individuals are increasingly commonplace. In most cases, these are vehicles with controls such as pedal operation, seat configuration etc. rearranged to suit the driver. However, when the impairments are vision related the accommodation is not as simple. In most interactions of the driver with the environment, the visual senses are the primary sensors, and vision impairments may result in substantial limitations or restrictions of performance. When working on assistive devices for these impairments, it is important that adequate testing be done in actual (on-road) conditions for any remedial evaluation.

Early investigations on the effect of in-vehicle devices focused on minimizing the effect of interface configurations and on structuring the information needs for specific tasks (Alm (1993) and Alm & Nilsson (1995)). Epidemiological methods have also been used in identifying safety effects from crash data in Wang, Knipling & Goodman (1996). Tijerina, Parmer & Goodman (1998) used commercially available navigation systems in a vehicle on a test track to evaluate the distraction caused by entry of destination data, and found that destination entry in commercially available navigation systems takes longer then the time required for the use of controls, the radio or a cell phone. Many of the studies rely simulators for assessing changes in behavior, and by inference, the risk, caused by the use of in-vehicle devices. Simulators range from rudimentary laboratory rigs to sophisticated immersive vehicle simulators. In spite of the substantial cost difference between the entry level simulators and the high end systems, it is not clear that any particular configuration is superior in inducing high fidelity responses from subjects.

In some cases, evaluations of driving performance have been conducted on test tracks with controlled (or no) traffic. Examples of this include navigation systems (Antin, 1993) and age and vision (Wood, 1999). Green (1999) discussed the limitations of simulation studies, and the extension of such studies to on-road settings, and concluded that the ability to perform on-road testing was limited because navigation tasks may be classified as “too risky".

Although simulator and test track studies may pose a lower risk as compared with on-road studies, the absence of test results from on-road conditions raises the bar for asserting the risk associated with these tasks, and the resultant ambiguity often leads to delaying or weakening legislation that should be enforced to improve driver safety. The legislative developments relating to use of cell phone while driving in the US can serve as a case study on this topic. In other cases, such as the testing of assistive devices for visually-impaired drivers, the need for on-road testing may be mandated by legislation. When on-road testing must be conducted, extreme care must be taken to anticipate and avoid exceptional conditions that may arise during testing. This paper describes some of the issues that must be considered in the design of on road test routes for driver testing. General design issues are considered and then special design elements are discussed for two specific studies: the impact of cognitive or other distractions (e.g. mobile phone use) on driving performance of normally sighted drivers and the effectiveness of vision devices as driving aids for the visually impaired.

## 2. METHODS OF DRIVER PERFORMANCE EVALUATION

On-road tests for assessing driver performance require tracking of one or more metrics related to the performance of the driver. The performance under "normal" or "un-distracted" conditions is compared with “assisted” or “distracted” driving conditions. The following is a categorization of the metrics typically used for driver performance measurement:

### *2.1 Vehicle Metrics*

These metrics relate to vehicle position and orientation, and include lane position, driving speed, vehicle steering angle, brake distances etc. Tracking metrics of this type requires vehicle instrumentation, although sufficiently accurate GPS systems can also be used instead.

### *2.2 Biological Metrics*

These are computed by tracking biological parameters such as heartbeat, blood pressure, eye-positions, pupil diameter, blinks etc. Improvements in instrumentation have now made it possible to track these metrics with portable devices in-situ. However the data analysis task remains formidable, particularly when tracking driver eye movements, where the eye position data has to be interpreted in conjunction with the scene video.

### *2.3 Cognitive Metrics*

These metrics relate to the locus of attention or cognitive workload of the driver. Estimates of these metrics are usually obtained through interaction with the driver to assess the performance of a driver under different conditions. Examples of such metrics include detection of traffic signs, time to perform a given computation, the accuracy in committing items to memory etc. Evaluating such metrics require interaction with the driver, usually by an accompanying investigator. Because of the obvious interference that the process of evaluation itself causes, these metrics should only be evaluated sporadically, with sufficient spacing between successive evaluations to maintain independence.

### *2.4 Subjective Metrics*

These include norms of driving performance that are difficult to assess by automated means, or do not include cognitive issues of driving workload. Interactions with other drivers, responses to pedestrians, obstacles etc. would be evaluated by such measures. Subjective metrics are usually assessed by observation, and scoring schemes must be designed carefully to ensure a uniform standard of evaluation is applied across different drivers and evaluators.

The classification above is not intended to be orthogonal - i.e. there is possibly significant correlation between the metrics in and across the categories. The primary issues concerning these metrics when designing a route are the sampling frequencies and the locations where they should be assessed. Biological metrics and vehicle metrics can be sampled continuously and require little to no intervention. However, because of interface issues, the maximum run length and the cumulative error may limit the duration of each segment of a drive, requiring an interruption to recalibrate and reset sensors. Cognitive and subjective metrics may require proper spacing because of the possible interaction of successive measurements.

## 3. DESIGNING TEST ROUTES FOR EVALUATING DRIVER PERFORMANCE

Routes designed with the goal of evaluating the impact of cognitive distractors on driving performance of "normal" or un-impaired drivers differ from routes designed to evaluate the effectiveness of vision devices as driving aids for the visually impaired. In the former case routes should be designed to take the driver through normal and exceptional circumstances in a controlled manner whilst collecting data to compare the metrics of interest against baseline or standard or normal levels. In the latter, the goal is to evaluate the performance of an impaired driver to compare his/her driving performance against an acceptable (normal) level of performance with and without visual aids. Regardless of the specific goal, several issues common to both route design processes are discussed below:

### *3.1 Route design preliminaries*

Any route design involves selection of starting, intermediate (where data is to be collected) and ending points. Data may be collected continuously from the start to the end - however, this can complicate subsequent processing, especially when video data is collected as well. Other factors involved in determining the length of period for which the data is collected include the spacing of tasks that are to be given to the subject. In general, each data-collection run starts with an instruction period, and often, a calibration period. Instrument calibrations can be time consuming, especially if the calibration has to be done in the test vehicle. The number of tasks is limited by the length of the route, and the need to replicate tasks under different road conditions. The use of Global Positioning Systems (GPS) and automatic tracking programs can greatly facilitate in the initial planning of the route.

### *3.2 Length of task durations*

For each task, the required length of the testing period is the sum of the time required for the instructions and time that the subject may require for responding. If the location where the instruction is delivered is kept the same, the time and approximate distance that a vehicle moves during this period will be the same. However, the task execution time can vary substantially: when response time is small it is approximately constant; but when response time is large it is variable. In general, the variability of the task execution time is related to the mean task response duration. Pilot runs can be made to determine the variance, and a limit (e.g. $3\sigma$) within which the task execution is expected. Major intersections, stop signs, school zones, restricted driving zones should be excluded from this limit.

### *3.3 Visibility for task presentation*

Tasks should be located on the test route with clear visibility of the oncoming roadway for the duration of the task. Thus, regions with tight turns or hills that present blind spots should be avoided unless specifically included for testing purposes.

### *3.4 Congestion and additional traffic interaction*

The congestion of a roadway may also significantly affect a driver’s behavior. Regions of congestion can be identified during path evaluation and can be used to compare the driver’s behavior under certain conditions with non-congested sections of roadway. If congestion is indeed a factor to be considered in route design, additional traffic interaction should be included. In some situations the time and direction of travel may need to be adjusted to place subjects on sections of roadway with similar traffic densities. The day of the week may also need to be taken into consideration in some locations.

### *3.5 Intersections and the need for stop signs and traffic lights*

Stop signs or traffic lights also need to be considered explicitly during the route-planning phase. If the purpose of evaluation relates to identifying differences in a driver’s behavior through different types of intersections, a few additional issues should be taken into account: first, if there is a certain level of congestion that needs to be present in the intersection, the time of day when the experiment is to be completed must be adjusted; second, if the driver needs to stop at a traffic light to test a task completion, some method of triggering the traffic light as the vehicle approaches must be negotiated with the appropriate roadway authorities; third, directions given to the subject describing where to go must be presented at a reasonable distance from the intersection to allow for sufficient execution time. Routines for coping with contingencies should be put in place prior to the study.

### *3.6 Directions to subject and investigators*

Directions given before the test can either fully describe the route or provide an overview as a method of familiarizing the subject with where they will be traveling. In situations where directions are only given once, the route must be simplified to where the subject can easily remember the path of travel. For directions given in-route, each step must be reduced to components simple enough for the driver to remember, and cues should be given to the accompanying investigator for prompting the driver. A standard method for communicating with the test driver during the drive can reduce driver anxiety and augment safety.

### 3.7 Error reduction

The errors encountered in an on-road driver performance evaluation can be divided into three categories based upon cause: driver errors, experimenter errors and recording errors. Driver errors occur when a driver fails to perform some task as directed. These cannot be eliminated but may be reduced by a clear presentation of directions. When an error does occur, corrective procedure must be in place. Options for data markers should also be provided to clearly identify the sections corresponding to the error. Experimenter error occurs when a task or direction is not presented as indented. Hesitation in presenting tasks may transfer uncertainty to the subject, reduce safety and interfere with the accuracy of the experimental results. Recording errors can be avoided through testing and planning. In general, more problems will be encountered with more complex recording equipment. Electronic devices need to be tested for interference with other devices, and for accuracy under the range of environmental conditions that they will be used. The procedures for starting, ending and calibrating recordings must be identified to provide clear markers in the data collected, any variations must be noted.

Other issues that should be considered include speed limits, signs, number of lanes, road conditions, subjects familiarity with roads, bridges or tunnels, road lights etc.

## 4. EVALUATING COGNITIVE DISTRACTIONS BY TRACKING EYE MOVEMENTS

When selecting routes for assessing the performance of drivers by tracking eye movements, an important issue that requires consideration is the road direction and time of day of the test. Once a possible route or set of routes has been identified, the angle of sunlight at different times of the day should be taken into account. The vehicle should be oriented to minimize the amount of incident infrared (IR) radiation inside the vehicle. If the IR levels are high, some eye trackers cannot accurately track the location of the pupil. Even with the use of an IR shield, reflections of IR off the vehicles internal surfaces can cause problems. Test trials should be completed during various periods of the day. When possible, roads with dense foliage should be selected to provide additional insulation from IR.

When evaluating distraction, the use of a driver’s own vehicle is recommended to reduce the additional processing load of driving in an unfamiliar vehicle. The influence of driver comfort should also be evaluated to ensure that the data collected is not exceptionally influenced by driver fatigue. As an example, if the level of accuracy required is high, head mounted eye tracking devices are necessary for measuring eye positions. However, these can limit the length of time that a subject can drive comfortably. Frequent stops and recalibration points must be arranged along the route to improve driver comfort and data collection success.

## 5. EFFECTIVENESS OF VISION DEVICES FOR VISUALLY-IMPAIRED PEOPLE

In this section issues related to the design of an on-road test route for investigation of driving performance of visually-impaired people will be discussed. The following impairments and assistive devices are considered:

1. People with reduced visual acuity, driving with and without bioptic telescopes;
2. People with hemianopic visual field loss, driving with and without peripheral prisms.

Design criteria ensure that the route will contain a representative range of normal driving tasks, as well as specific tasks that are expected to be difficult for people with each type of vision impairment, and will be sensitive to evaluating performance with and without each vision device. The aim is to have design criteria that can be easily implemented to produce multiple versions of a route (for assessments with and without a device), which are as similar as possible and can be implemented at several study locations. This approach enables the use of the same route design for more than one type of vision impairment.

### *5.1 Type of vision impairment and relation to route design*

Difficulties encountered by visually-impaired drivers are expected to vary according to the type of vision impairment. Patients with reduced visual acuity have difficulty seeing details, for instance reading road signs and seeing objects at a long distance, however they usually have normal peripheral vision. In some jurisdictions they are permitted to use bioptic telescopes when driving (e.g. in 34 states in the USA: Peli, 2002). The bioptic telescope provides magnification to enable small details to be seen at a normal approach distance, and is used for reading road signs and scanning ahead for hazards (Kelleher et al., 1971; Corn et al., 1990, Peli et al., 2003). Specific sign reading tasks are therefore included in the route design.

By comparison, people with hemianopic field loss have normal visual acuity, but a restricted visual field (they lack half of the field on one side in both eyes). As they have a blind (non-seeing) hemi-field, they might miss objects in this field unless they scan effectively while driving. We planned to evaluate the effectiveness of peripheral prisms as a driving aid for people with hemianopia. Peripheral prisms provide expansion of the field into the blind hemi-field, thus increasing the probability of detection of objects on the blind side (Peli, 2000). As the blind hemi-field may be on the right or the left, an equal number of right- and left-sided tasks are included in the route design. Hemianopes may exhibit unsteadiness of steering (poor lane control) and difficulties with correct lane positioning; these aspects of driving performance are therefore incorporated in the design (Szlyk, 1993; Tant, 2002).

### *5.2 Route design specifics*

The route design comprises a pre-test section (about 10 to 15 minutes) and a scored test section of about 30 to 45 minutes on-road driving. The purpose of the pre-test section is to enable the subject to become familiar with the car (while driving around a parking lot or quiet roads) and for the driving instructor to ensure that the subject demonstrates adequate vehicle control and driving before proceeding to the scored section of the route. The scored section contains a variety of types of road including single and dual carriageway and motorways (highways), and intersections with and without traffic lights, stop and yield signs and roundabouts. The roads include sections with well and poorly marked edges, straight sections and bends, quiet and moderate traffic density.

The following scored tasks are included in the test section, with about one scored task every minute:

- Lane control: straight sections of road to evaluate steering and lane position control.
- Lane changing: on multi-lane road, to evaluate lane changes to left and right
- Curve taking: equal number of right and left bends to evaluate steering in curve taking
- Turns: equal number of right and left turns at intersections with and without traffic lights, stop and yield signs
- Crossing intersections: crossing intersections with and without stop and yield signs
- Roundabout: entering, driving round and exiting roundabout
- Parking: move forwards into and reverse out of designated parking space
- Reading signage: call out road signs seen along specific sections of road and find specific streets by street name signs
- Motorway (highway) driving: merging, exiting, and overtaking (also including one section of lane control, reading signage and lane changing).

The design also includes at least two locations where it is possible to pull off the road (to enable completion of scoring, if necessary) and planned "recovery" routes so that if a subject takes a wrong turning, minimal time will be taken in returning to the original route. In addition the route is designed so that there will be sufficient time for driver instructions to be given prior to each maneuver and driver response or action to be taken safely.

### *5.3 Assessment procedures*

Based on route design and the aspects of driving performance to be evaluated, a template score sheet was developed for each of the main scored driving tasks (i.e. turn, change lane etc). Later, these are then made route specific by the addition of road and intersection details, and compiled into a booklet according to the sequence of driving tasks along the route.

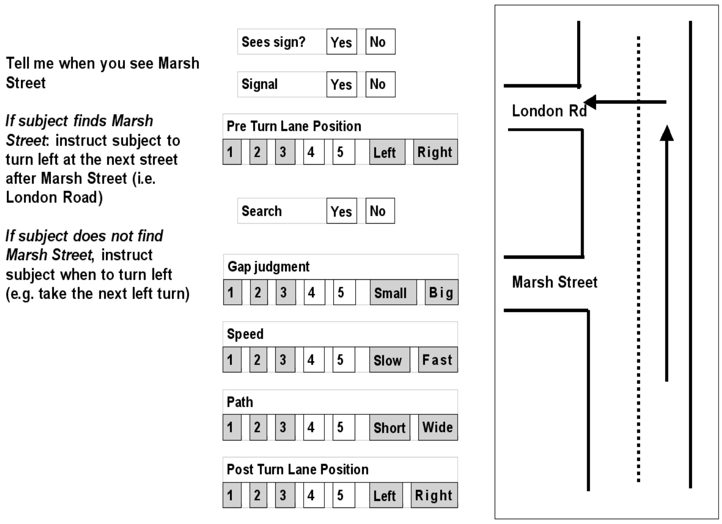


**Figure 1: Example of a task score sheet (find a specific street and complete a left turn at the following street).**

Driving performance is scored using a series of items listed in the middle column (Figure 1), while instructions to be read to the driver are in the first column and a depiction of the road layout is given in the last column. The specific driving skills (items) to be scored for the maneuvers are: lane position, gap judgment (the gap or distance between the test vehicle and vehicles from one or both sides when crossing or moving into their lane), speed (whether speed is appropriate for the maneuver), path (the path taken when going around a curve), steer steadiness, spacing (following distance) and speed match (when merging or overtaking). Each of these items is scored on a five-point scale. Scores of 1-3 represent various levels of unsatisfactory performance (from 1, driving evaluator had to take control of car, to 3 unsatisfactory but does not compromise safety), while 4 and 5 represent increasing levels of satisfactory performance. To aid completion of a score sheet, the boxes for unsatisfactory scores (1, 2 and 3) are shaded grey (Figure 1). An additional score is also required when an unsatisfactory mark is given. This score is based on a binary choice, recorded in the boxes to the right of the 5-point scale, indicating "excesses" in performance e.g. if the lane position is too far to the left/right, gap judgment is too small/too big (Figure 1). In addition other items are scored Yes/No, such as whether the signal indicator is used (signal), whether the subject scans before performing a maneuver (search), whether the subject obeys traffic signals or reads traffic signs.

In addition to the task score sheets, a "global" score sheet was developed to evaluate overall driving performance along the whole route (not just the scored tasks). This evaluation is performed at the end of the test ride and includes interaction with other traffic, vehicle control, anticipatory skills, adjustment of speed to traffic conditions, reaction to unexpected events and an overall rating of driving performance, including the question "Would you give this person a license?".

Two evaluators travel in the dual control vehicle. The primary evaluator, who sits in the front passenger seat, gives instructions to the driver and is responsible for safety (taking control of the car if necessary). He/she only completes the global score sheet. The second evaluator, from the back seat, completes each task score sheet (immediately after each task is performed) and the global score sheet.

## 6 DISCUSSION AND CONCLUSIONS

This paper discusses the design of test routes for evaluating driver performance. Two different types of on-road driving evaluations have been presented - the first for assessing the impact of distractions on drivers, and the second for evaluating the driving capability of drivers with and without vision assisting devices. The designs for these two on-road evaluations were set up independently by two groups of investigators. In addition to the identification of common features involved in the route design process for both evaluations, several noteworthy lessons emerged from the comparison: 1. When objective methods such as eye movement tracking are used for assessing driver performance, care must be taken to ensure that the data collection process is robust. This can be attained by proper vehicle preparation (instrumentation), automation of task presentation and data recording, calibration and route optimization. Objective methods are more suitable for evaluating the performance of normal drivers, and can be used to help identify exceptional behavior under test conditions. 2. Subjective scoring methods should be used for assessing performance when the number of metrics to be evaluated is large, and an integrated evaluation is desired over specific route sections. These methods are more suitable for evaluating the performance of impaired drivers under normal or usual driving conditions. However, better data may be obtained by integrating features from both methods. When subjective scoring methods are used, it may be possible to automate the collection of vehicle metrics relieving the observer from this monitoring activity. Similarly, when using objective methods such as eye movement tracking, better information about the driver's performance may be recorded by including subjective metrics vis-à-vis scoring sheets. Given the increasing intrusion of in-vehicle devices, and a growing population of drivers with impairments, the need for on road testing is likely to intensify. The guidelines suggested in this paper could reduce the testing effort and risk.

## REFERENCES

Alm, H. & Nilsson, L. (1995). The effects of a mobile telephone task on driver behavior in car following situation. Accident Analysis and Prevention, 27(5), 707-715.

Alm, H. (1993). Route navigation. deciding driving information needs. In S. Franzen A. Parkes, Driving future vehicles (187-192). London: Taylor and Francis.

Antin, J. (1993). Informational aspects of car design: Navigation. In B. Peacock W. Karwowski, Automotive Ergonomics ( 321-337). Washington DC: Taylor and Francis.

Corn AL, Lippmann O & Lewis, M.C. (1990). Licensed drivers with bioptic telescopic spectacles: user profile and perceptions. RE:view. ( 21). 221-230.

Green, P. (1999). Visual and task demands of driver information systems.(Report UMTRI-98-16). Ann Arbor, MI: University of Michigan Transportation Research Institute.

Kelleher DK, Mehr EB & Hirsch MJ. (1971). Motor vehicle operation by a patient with low vision: a case report. Am J Optom Arch Am Acad Optom. 8, 773-776.

Peli E, Bowers AR, & Apfelbaum D. (2003) Driving habits and experiences of bioptic drivers: a cross-sectional survey. Presented at Vision In Vehicles 10, Granada, Spain.

Peli E. (2002). Low vision Driving in the USA: who, where, when, and why. CE Optometry. 5, 54-58.

Peli, E. (2000). Field expansion for homonymous hemianopia by optically induced peripheral exotropia. Optom Vision Sci. 77, 453-464.

Szlyk JP, Brigall M & Seiple W. (1993). Effects of age and hemianopic visual field loss on driving. Optom Vision Sci., 70, 1031-1037.

Tant MLM (2002). Visual performance in homonymous hemianopia: assessment, training and driving. PhD Thesis, University of Groningen, Netherlands.

Tijerina, L., Parmer, E. & Goodman, M. (1998). Driver workload assessment of route guidance systems destination entry while driving: a test track study. In Proceedings of the 5th its world congress ( October). Seoul, Korea.

Wang, J., Knipling, R. & Goodman, M. (1996). The role of driver inattention in crashes: new statistics from the 1995 crashworthiness data system. In Proceedings of 40th annual meeting of the association for the advancement of automotive medicine ( October). Vancouver.

Wood, J. M. (1999). How do visual status and age impact on driving performance as measured on a closed circuit driving track? Ophthalmic and Physiological Optics 19(1): 34-40.