Draft Recommended Practice - SAE J-2396

Revised 12-98
(Not in SAE document format)


Forward: This document is intended to be as consistent with the International Organization for Standardization (ISO) 15007 as possible and is developed to extend that document to include factors which affect the results of video based visual allocation (VBVA) studies.

Contents:
I. Introduction
II. Video Based Eye Glance Capture Techniques
III. Scope
IV. Field of Application
V. Normative reference(s)
VI. Measuring Eye Glance Behavior in Driving
VII. Definitions
VIII. Developing a Glance Allocation Measure Data Base
IX. Specification of Independent Variables
   A. Target Factors
   B. Driver Factors
   C. Experimental Conditions
X. Testing the Validity of Glance Data
I. Introduction

Vision provides the primary source of information available to the driver. Information is gathered by looking at objects and events which in turn enables decision making, control, and navigation of the vehicle in the road traffic environment. Assessment of driver visual behavior (eye glance behavior) provides a method to quantify the driver’s visual allocation to the roadway, in-vehicle information sources, controls and mirrors and as such can be a useful tool in many ergonomic studies of the driver. These would include evaluation of the visual demands of in-vehicle information devices, assessment of visual distraction, driver fatigue, workload, individual differences (e.g., novice vs. expert) and basic studies of visual allocation in driving.

Transport Information and Control Systems (TICS) and other advanced display and control systems associated with Intelligent Transportation Systems (ITS) can present a range of driver related information. If these visual displays have associated controls (e.g., to select a zoom level or menu option), then the associated hand-controlled activities may also be visually guided and become part of the visual demand associated with a display application. For this reason, it may be important to consider not only the visual behavior in relation to information display, but also the duration and frequency of glances associated with driver control actions.

In the past, comparisons between separate evaluations of specific vehicle systems in different test environments have been made more difficult by dissimilar approaches in experimental techniques, operational definitions and analysis methods.

This recommended practice has been developed to give guidance on the terms and measures relating to the collection and analysis of driver visual behavior data. This approach aims to assess how drivers respond to vehicle and equipment design, the road environment, or other driver related tasks in both real and simulated road conditions. It is based on the assumption that efficient processing of visual information is essential to the safe performance of the driving task in a given driving situation.

Practical assessments of drivers in real or simulated environments are conducted to quantify the allocation of visual behavior to specified targets. It may be quantified by the location, duration, and frequency of glances to a specified target in the visual scene. This approach often employs commonly available video-recording equipment. However, it does not preclude the use of more sophisticated technologies such as eye movement techniques, which may elicit additional driver visual behavior information.
II. Video Based Eye Glance Capture Techniques

In order to better to understand the definitions to follow it is important to discuss the nature of the video data collection techniques in use. Equipment for data acquisition of driver visual allocation using video techniques can range from simple, portable, low cost systems to more sophisticated, high capacity and more expensive systems. These systems are for use on the road or for simulators and can be used with or without the presence of an on board experimenter. Although other data can be developed from video systems which may relate to glance behavior, (e.g., hand location on steering wheel, traffic density and headway estimation (distance between the front bumper of the test vehicle and the rear bumper of the lead vehicle)) the emphasis of this recommended practice is on measurement of glance behavior.

A. Simple Systems

A simple system to answer a very specific question, (e.g., glance frequency and glance durations to an experimental instrument panel display), would at a minimum have the following components:

1) an inverter for 120 volt power (road use only)
2) a time code generator
3) two cameras (either a miniature single chip complementary metal oxide semiconductor (CMOS) black and white (b/w) camera or a b/w infra red sensitive surveillance camera for day and night use) with one camera directed at the driver’s eyes and the other on the road.
4) video monitor for use by an experimenter
5) four channel video multiplexer (for more than one camera)
6) video cassette recorder (VCR) for data collection
7) infra red light source for night operation

This system would be portable, easy to assemble, require no special software, and be operational in a few weeks with off the shelf hardware costing about $2,000 (1998 U.S. dollars). It would of necessity require more effort in data reduction and be subject to timing interpretation errors.

B. Advanced Systems

A more sophisticated system makes use of several miniature cameras with digitization and compression of the camera output. The system is high capacity, tapeless and requires no time code generator. Such a system can record for long periods without experimenter intervention and is more flexible than a VCR in capturing video clips of time critical events. Such a system requires more time to assemble, some complimentary software and has a component parts cost of some $5,000 (1998 U.S. dollars).

The following components would be required:

1) power supply - 12v to 120v, and 120v to +/- (5v and 12v)
2) cameras (miniature single chip CMOS b/w)
3) computer monitor, keyboard and mouse
4) on-board computer complete with removable hard drives, mounted in a frame and carrier system, moving picture expert group - the coding of audio and visual information in a compressed format (MPEG) video encoder and RS-232 interfaces
5) other analog/digital (A/D) and digital/digital (D/D) connections as required by other sensors, (e.g., headway, lane position, gas pedal position etc.)
6) infra red light source for night operation

III. Scope

This recommended practice defines key terms and metrics applied in the analysis of driver eye glance behavior. It can be applied in environments from real world trials to laboratory based driving simulator studies.

IV. Field of Application

The procedures described in this recommended practice could also apply to more general assessments of driver visual behavior in the absence of TICS or other advanced display and control systems associated with ITS. Driver workload studies, modality interference from use of cell phones, mirror redesign, situational awareness, and the effects of driver stress from sleep loss and trip delays are just a few of the studies that would benefit from a standard practice for measuring visual allocation. The metrics and definitions described below are intended to assist development of a common source of reference for driver visual behavior data. Data collated and analyzed from this recommended practice allows comparisons to be performed across different device applications and experimental scenarios. It should be noted that the following definitions and measures would also apply to eye movement techniques.

V. Normative reference(s)


Before defining terms it is important to understand the process of visual perception through saccadic eye movements. The main function of a saccade is to change the point of fixation to direct the most sensitive region of the retina (the fovea) to a particular object of perception (Yarbus 1967). These saccades are typically less then 20 degrees in amplitude and have a duration of .01 to .02 seconds. After one or more saccades, fixations of .2 to .4 seconds are made to process information. A glance is considered as a series of fixations at a target area until the eye is directed at a new area. Operationally, the glance duration includes the prior transition time (to be consistent with ISO 15007). These transition times typically range from .10 to .5 seconds depending upon the distance between the two targets. We can also define dwell time to be the glance duration minus the prior transition time or the sum of all fixations and saccades in the target area between transitions or shifts of gaze to other target areas.

With sophisticated eye movement systems it is possible to measure the duration of fixations and transition times. However, for video based systems accuracy is limited to static target glance durations (usually between .5 to 3+ seconds). It should be noted that gazes to the road usually involve several glances. A graphical depiction of the visual allocation process would help describe the definitions to follow. See Figure A.1.

The chronological relationship of driver visual allocation between and among targets is illustrated in Figure A.2. Each horizontal segment represents a fixation (alignment of the eyes so that an image of the fixated target falls on the fovea of the eye for a given time period). Slanted lines represent saccades or transitions.

It is also important to review the basic data reduction protocol for VBVA studies. Data reduction is, of course, dependent upon the data collection system employed. In the simplest form, video frames are indexed forward on a VCR and the frame count number recorded manually when the eye begins and leaves a specified target. Typically each frame, a single picture based upon the speed of the camera, constitutes 1/30 second for U.S. VCR systems. Programs can be written to do the subtraction and conversion of frames to seconds. In other systems involving a time code generator and special computers (and software) the data reducer stops the advance of the frame and timing and location data go directly to a spreadsheet for analysis. Special timing and interpreting software permits keyboard control of location events (with time automatically recorded and synchronized with concurrent engineering measures such as lane position, travel speed and headway).

Clearly there are four strategies possible for defining a glance: (See Figure A.2)
1) ignore transition times
2) assign transition times forward to the next target location
3) assign transition times back to the previous target location
4) add both transitions to and from a target to that target
To be consistent with ISO 15007, it is recommended that the glance duration to target A include both the transition from B to A plus the fixation time at A (T1+ F2 in Figure A.2) using the second strategy above. Any glance duration less than .2 seconds should be ignored and added into the following target.

VII. Definitions

7.1 direction of gaze - the target or target location to which the eyes are directed
7.2 dwell time – the sum all fixations and saccades within the target area between transitions or shifts of gaze to other target areas
7.3 fixation - alignment of the eyes so that the image of the fixated target falls on the fovea for a given time period
7.4 frame – the basic unit of observation for data reduction based on the video data capture rate. It is one of the successive pictures recorded on a video tape or on digitized video. The data reducer (or data reduction software) examines a video display of the driver’s eyes frame by frame, to determine the driver eye fixation location.
7.5 glance duration - the time from the moment at which the direction of gaze moves toward a target (e.g., the interior mirror) to the moment it moves away from it. This includes the transition time to that target (See Figure A.2). A single glance duration may also be simply referred to as a glance.
7.6 glance frequency - the number of glances to a target within a pre-defined sample time period, or during a pre-defined task, where each glance is separated by at least one glance to a different target.
7.7 glance probability - the probability of a glance to a given location. This may be defined, for location A, as the sum of all transitions to location A divided by the sum of all transitions between all pairs of locations in the sample interval. Note that this does not reflect the proportion of glance time on location A, but instead reflects the proportion of transitions to location A.
7.8 link value probability - the probability of a glance transition between two different locations. Operationally, the link value probability between target locations A and B is defined as the number of glance transitions from A to B plus the number of glance transitions from B to A; this sum divided by the total number of glance transitions between all pairs of locations in the sample interval.
7.9 saccade - the brief movement of the eyes between fixations. See Figure A.2.
7.10 sample interval (period) - a reference time period that constitutes a sample of interest (e.g., an in-vehicle task or maneuver) in the video data. Usually, this will be the time associated with a reference event.
7.11 separation angle (in-vehicle devices) - the angle subtended at the eye between the center of two in-vehicle targets. This angle should be > 20° unless a sampling to such targets is commanded or if calibration demonstrates better discrimination.
7.12 target (target location) - a pre-determined area within the visual scene, (e.g., a rear view mirror). See Figure A.1. For commanded visual tasks it can be a specific instrument.
7.13 *time off road scene* - the total time between two successive glances to the road scene which are separated by glances to non road targets

7.14 *transition* – a change in eye fixation location from one defined target location to a different target location

7.15 *transition time* - the duration between the end of the fixation on a target location and the start of the fixation on another target location. See Figure A.2.

7.16 *TIC or ITS device* - a device used to present information or a control which requires visual fixations to locate and operate

7.17 *visual angle (for non road scene targets)* - the angle subtended at the eye by the largest dimension of a viewed object

### VIII. Developing a Glance Allocation Measure Data Base

#### A. **Required Statistical Measures**

Over a sample period of interest (experimental condition) for each subject and subject grouping, compute glance frequency and glance duration for:

1) Road scene ahead  
2) Left side mirror  
3) Right side mirror  
4) Center vision mirror  
5) In-vehicle device(s)  
6) Instrument panel (IP) (radio or gauges of interest)  
7) Time off road scene (this would be derived by measuring the time allocated to non road scene targets between glances to the road scene)  
8) Out of view - the eye located on a target not specified or out of camera view  
9) Number of over the shoulder head turns.

For the above measures specify:

a) N (the number of observations)  
b) arithmetic mean  
c) variance  
d) distribution (for distributions of glance durations, use .1 second class intervals).

In some applications data conditioning may be necessary to do transformations to handle outliers, (e.g., log normal). If the accuracy of the data permits, similar calculations could be made for transition times. Other derived measures of interest might include number of glance durations above 2.5 seconds (a recognized measure of safety), glance probability, link value probability and total time off road scene.

The measures derived above for a given set of experimental conditions should be compared with baseline conditions, (e.g., without the use of an experimental device).
B. **Optional Statistical Measures**

Depending upon computer statistical packages available, the following statistics can provide greater insights from the data above.

1) 95% confidence limits
2) quantiles
3) medians
4) 3rd and 4th moments
5) cumulative distributions
6) box plots with outliers
7) coefficient of variation
8) correlation with other experimental measures, (e.g., speed, lane exceedances, age etc.)

IX. **Specification of Independent Variables**

In order to allow comparison of similar studies by different experimenters, it is critical to catalogue the conditions of the study, namely the unique set of subjects, road conditions and targets (devices) to be evaluated. Moreover, it is also of interest to test how eye glance data is affected by these independent variables. While this specification is associated with any good human factors study, adoption of this recommended practice by researchers and evaluators will insure proper interpretation of experimental data.

A. **Specification of Target Factors**

1. Target Classification

Typically targets include road scene, mirrors (left, center, and right side), in-vehicle displays, controls and typical IP devices, (e.g., speedometer). Depending upon the care exercised in calibration (see below) adjacent instrument panel targets can be differentiated depending upon their separation angle. It is expected that a separation angle of at least 20° is needed to differentiate visual sampling between adjacent targets unless calibration tests demonstrate differentiation of adjacent targets at smaller angles.

All non road scene targets should be specified by:
   a) name
   b) the location of the center of each target nominally measured horizontally and vertically from the eye when the latter is directed at the road ahead

2. TIC or Information Display Specification

Specify:
   a) overall size of display – height and width in centimeters
   b) size of alpha numerics and symbols in millimeters (include font type). A photograph of the display(s) with a reference object included will document the measurement.
c) color and contrast of alpha numerics and symbols with background - demonstrate with color photos.
d) any special display characteristics of importance to the evaluation of the data - flashing messages, refresh rates, menu choices, error recovery procedures, system latencies, data entry syntax, etc.

3. Target Calibration Procedures (studies with on-board experimenters)
Prior to data collection and at intervals within the test runs (usually every 30 minutes) calibration exercises must be performed to account for subject postural changes during data collection and to aid in data reduction. The subject is asked to visually fixate on each of the targets while being videotaped. This calibration tape segment is used to assist data reducers in making decisions on where the driver is fixating with any particular glance.

B. Specification of Driver Factors

1. Subject characteristics
For all subject drivers provide the following information:
   a) driver age (or age brackets)
b) gender
c) licenses held
d) corrected far visual acuity (use Snellen chart)
e) corrected 28 inch visual acuity (IP distance) – use modified Snellen chart
f) driving experience (years)
g) experience with TICS (number of practice trials)
h) experience with the test vehicle (minutes)
i) special skills required - manual dexterity, voice activation requirements, etc.

2. Subject Instructions
Provide specific instructions in writing to each test subject. These should include the following:
   a) alleged purpose of the testing
   b) specific maneuvers required (e.g., car following, headway maintenance, lane placement, speed maintenance)
c) required use of in-vehicle devices
d) safety instructions (priority of safety over data collection)
e) test termination options open to subject at their request

3. Subject Debriefing
Following each experimental session the subjects will be debriefed to include the following:
   a) subject perception of the real purpose of the study
   b) perceived realism of the tests
c) problems in carrying out the instructions
d) awareness of being videotaped
e) subjective evaluation of the TICS (display devices, including suggestions for redesign)
f) comfort level with the test vehicle

C. Specification of Experimental Conditions

1. Road Specification
If the road tests are divided into different segments associated with test trial time periods, then for each segment specify:
   a) number of lanes (each way)
   b) speed limits
   c) limited access (if relevant)
   d) presence of vertical and horizontal geometry and a qualitative estimate of the degree of curvature of hills and curves during data collection

2. Traffic Specification
For each segment:
   a) estimate from the videotape using a calibrated lead vehicle of known width the average headway (meters) between the front bumper of the test car and the rear bumper of the lead vehicle only when subjects are required to car follow.
   b) estimate the traffic density (light, medium, heavy). Using 2 seconds or less as a time headway (headway/speed) for car following, consider light traffic to be car following 0-20% of the trial time, medium traffic 21-60% of the trial time and heavy traffic as 61-100% of the trial time.

3. Lighting Specification
For each test segment specify:
   a) the overall light level (day, dusk, dark)
   b) whether overhead lights were present at night
   c) headlight beam pattern (low vs. high) at night
   d) if the sun angle to the driver (if dusk or dawn) affected visual search

4. Vehicle Specification (road vehicle or simulator)
Specify:
   a) size of vehicle (e.g., full size, compact, truck, etc.)
   b) type of transmission (automatic vs. stick shift)
   c) for a simulator specify screen size and any special characteristics that might affect visual search
X. Testing the Validity of Glance Data

While variations in driver glance behavior detection will differ by subjects and experimental conditions, the validity of the results can be tested against the literature (see references). For example, left mirror detection sampling, common to all driving tasks, will have glance durations ranging from .8 to 1.3 seconds. Provided no unusual instructions were presented to the subject relative to left mirror use, glance data for left mirror sampling beyond this range should be rechecked to ascertain any errors in data collection or reduction.