Sequential In-Vehicle Glance Distributions: An Alternative Approach for Analyzing Glance Data

Yusuke Yamani, William J. Horrey, Yulan Liang and Donald L. Fisher

DOI: 10.1177/0018720814560225

The online version of this article can be found at:
http://hfs.sagepub.com/content/early/2014/11/13/0018720814560225
Sequential In-Vehicle Glance Distributions: An Alternative Approach for Analyzing Glance Data

Yusuke Yamani, Old Dominion University, Norfolk, Virginia, William J. Horrey, Yulan Liang, Liberty Mutual Research Institute for Safety, Hopkinton, Massachusetts, and Donald L. Fisher, University of Massachusetts Amherst, Amherst, Massachusetts

Objective: The aim of this study was to illustrate how a consideration of glance sequences to in-vehicle tasks and their associated distributions can be informative.

Background: The rapid growth in the number of nomadic technologies and in-vehicle devices has the potential to create complex, visually intensive tasks for drivers that may incur long in-vehicle glances. Such glances place drivers at increased risk of a motor vehicle crash.

Method: We used eye-glance data from a study of distraction training programs to examine the change in glance duration distributions across consecutive glances during the performance of various in-vehicle tasks.

Results: The sequential analysis across trained and untrained drivers showed that the proportion of late-sequence glances longer than a 2-s threshold among untrained drivers was almost double the number of such glances for the trained drivers, that the third and later glances were particularly problematic, and that training reduced the proportion of early- and later-sequence glances.

Conclusion: Examining how the duration of off-road glances varies as a function of their order in a sequence of glances and the visual demands of the task can offer important insights into the change in the distracting potential of in-vehicle tasks across glances and the effects of training.

Application: The sequential analysis of in-vehicle glance data can be useful for researchers and practitioners and has implications for the development and evaluation of training programs as well as for task and interface design.

Keywords: driving safety, attention maintenance

INTRODUCTION

Drivers often engage in in-vehicle tasks that can distract them from active monitoring of the forward roadway, leaving them exposed to potential threats. Studies have found that especially long off-road glances, including those to in-vehicle and nomadic devices, elevate the risk of a crash (e.g., Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006; Wierwille & Tijerina, 1998). Complex in-vehicle tasks requiring a greater number of especially long glances to complete are particularly problematic for driving performance and road safety (e.g., Green, 2002; Horrey, 2009; Victor, Harbluk, & Engström, 2005).

In light of these outcomes, authors of several studies have argued for a shift from the reliance on measures of central tendency toward the examination of glance distributions, particularly at the safety-critical tail end (e.g., Horrey & Wickens, 2007; Pollatsek, Divekar, & Fisher, 2013). However, just as summary measures of glance durations could fail to reveal important information about distributions of glances (Horrey & Wickens, 2007), so could aggregate distributions across all glances fail to reveal important information about the distribution of each successive in-vehicle glance.

Models of supervisory control posit that eye movements reflect the drive to reduce uncertainty regarding the state of the world at different locations, such as the road ahead or instrument panel (e.g., Carbonell, Ward, & Senders, 1968; Senders, 1983). The duration of each glance to a particular location reflects the amount of processing that is required or the relative difficulty in resolving information needed to reduce uncertainty (e.g., Henderson, 2003). It follows that glance duration could vary within a task as a function of the information that is obtained.
during each glance. As such, aggregating or collapsing glance duration data across entire task interactions or across tasks themselves could mask important differences in the nature of glances across each successive glance. For example, Lee, Roberts, Hoffman, and Angell (2012) found that under certain task conditions, glances occurring later in a given task were longer than earlier glances, and the frequency of especially long glances increased. Liang, Lee, and Yekhshatyan (2012) also found that the length of the current glance was the most sensitive indicator of crash risk relative to models that considered aggregated measures of glance history.

This brief report aims to illustrate the utility of an examination of empirical probability distributions of glance duration for each glance in a sequence of glances during various types of in-vehicle tasks. It draws upon data gathered in a previous driving simulator study designed to reduce the proportion of especially long in-vehicle glances (Divekar et al., 2013); specifically, researchers developed a training program, FOrward Concentration and Attention Learning (FOCAL), aimed at training novice drivers to glance inside the vehicle for no longer than 2 s (e.g., Klauer et al., 2006). The training has been shown to be effective in reducing especially long glance durations in different settings (Divekar et al., 2013; Pradhan et al., 2011). However, the distribution of glance durations for a given task was obtained by aggregating the data across all glances required to complete each of seven different tasks. It follows from the discussion earlier that analysis of aggregated data could mask important nuances in the underlying glance behavior. For example, it is not known whether the training is as effective for later glances in a sequence as it is for earlier glances. Such an outcome could hold important ramifications for understanding of such training and in the coordination of sequential eye glances. Nor is it known whether the sequence of glance durations varies across different tasks or is only a function of the order of a glance in a sequence. Although the current data are derived from a training study, we argue that this approach could offer important insight in other applications as well, such as interface design and evaluation.

METHOD

In the Divekar et al. (2013) study, 40 young drivers (23 males and 17 females; $M = 16.5$ years, range = 16–18 years) with a junior operator’s license were randomly assigned to a FOCAL-trained or a placebo group. The simulator consisted of a full-sized Saturn sedan with three screens subtending 135° horizontally (Realtime Technologies, Royal Oak, MI). The driving environments were populated with vegetation, objects, and randomly parked vehicles on two-lane or four-lane roadways. While driving, participants performed tasks inside and outside of the vehicle; the current effort focused on the seven in-vehicle tasks (described in Table 1). Auditory instructions and alerts indicated initiation and termination (after 15 s) of the tasks, respectively, and each task was initiated at a fixed location in each scenario. The location and timing of the in-vehicle tasks were controlled across drivers, and in each instance, drivers had 15 s to complete the task. Eye position was sampled at 25 Hz with a head-mounted eye tracker (Applied Science Laboratories, Bedford, MA). The postprocessed output was combined with the video of the driving scene with a superimposed crosshair indicating the position of the eye in each frame.

RESULTS AND DISCUSSION

The duration of an off-road glance was defined as the time elapsed between the moment when a driver’s eyes transition away from the forward roadway up to and including the transition where they return (after Pradhan et al., 2011; Yamani, Samuel, & Fisher, 2014; note that the ISO 15007 standard does not include the final transition). In total, 767 glances for the FOCAL-trained group ($M = 1.32$ s, $SD = 0.88$, range = .04–10.11) and 735 glances for the placebo condition ($M = 1.65$ s, $SD = 1.11$, range = .04–10.44) contributed to the analysis. Averaging across all tasks and all glances in a task sequence, Divekar et al. (2013) showed that FOCAL training decreased the proportion of glances greater than various threshold values. For the current analysis, focused on training effects, glances were divided into 500-ms bins based on their order in a task sequence for both
### TABLE 1: Task Descriptions, Subtasks, and Visual Demand Coding for the Various In-Vehicle Tasks Performed by Drivers

<table>
<thead>
<tr>
<th>Task</th>
<th>Description and Subtasks</th>
<th>Visual Demand</th>
</tr>
</thead>
</table>
| 1. High beam | Engage the high beams  
(1) Locate the control  
(2) Engage the high beam  
(3) Verify | 1 |
| 2. Hazard lights | Engage the hazard lights  
(1) Locate the control  
(2) Engage the hazard lights  
(3) Verify | 1 |
| 3. Climate control | Adjust the temperature (air conditioning) controls to a specific setting  
(1) Locate the control  
(2) Adjust the temperature controls  
(3) Verify settings  
Repeat (2) and (3) until completion | 2 |
| 4. CD search | Find a specific CD from a case containing 24 CDs  
(1) Locate the CD case  
(2) Reach and get the case  
(3) Open the case  
(4) Inspect a CD  
Repeat (4) until completion | 3 |
| 5. Map search | Find a specific street on a laminated map  
(1) Locate the map  
(2) Get the map  
(3) Inspect a street name  
Repeat (3) until completion | 4 |
| 6. Radio tuning | Tune the radio to a specific frequency in either AM or FM  
(1) Locate the control for band  
(2) Choose band (AM or FM)  
(3) Locate the frequency control  
(4) Adjust frequency  
(5) Verify frequency  
Repeat (4) and (5) until completion | 2 |
| 7. Coin search | Find a specified amount of spare change from a box located near the glove compartment  
(1) Locate the basket of coins  
(2) Draw a coin  
(3) Inspect the coin  
(4) Do mental summation  
(5) Place the coin  
Repeat (3) through (5) until completion | 1 |

*Note. Visual demands were rated by subject matter experts (three of the current authors) following a method prescribed by Wickens (2002; Horrey & Wickens, 2003). The table values represent the aggregated scores (rounded to nearest integer). For each subtask, visual demands were coded according to lowest ordinal algorithm to rank demands. For example, 0 is assigned for a subtask that requires no visual demand, 1 involves little visual demand, 2 requires more, and so on. The net result of this exercise was a visual demand vector that could characterize the visual demands of the task at early, middle, and late stages of its execution.*
the FOCAL-trained and placebo-trained conditions. We limited our sequential analysis to the first seven glances (mean number of glances in the tasks ranged from 2.7 to 7.6 in the placebo-trained group and from 4.7 to 6.9 in the FOCAL-trained group). Figure 1 illustrates the complementary cumulative distribution function (CDF), $1 - F(t)$, for each of the seven sequenced glances, which are collapsed across the in-vehicle tasks (inspection of the sequential glance distributions for each task showed them to be similar). $F(t)$ is the CDF, which for each value of $t$ represents the probability that a glance will have a duration less than or equal to $t$. Thus, the complementary CDF of in-vehicle glance durations indicates the probability that the glance duration will be greater than or equal to $t$. Visual inspection reveals more short glances for the first and second glances in the sequence and less frequent but roughly equal numbers of short glances for later glances in the sequence (Glances 3 through 7; Figure 1).

We conducted a series of the Kolmogorov-Smirnov (K-S) tests with Bonferroni correction to examine statistically whether samples of glance durations for each pair of glance sequences arose from the same population distribution. The results of these tests corroborate the results from the visual inspection and strongly suggest that the distribution of glance durations for all glances in the sequence from the third glance onward are likely to have arisen from the same distribution, producing similar shapes of the distributions. It is equally clear, and of interest, that the training effect on the later glances in the sequence is as strong as earlier glances in the sequences. For example, the proportion of first glances longer than 2 s is 0.07 for the FOCAL-trained group and 0.16 for the placebo-trained group, a difference of 0.09. The proportion of glances later in the sequence (3 through 7) is 0.20 for the FOCAL-trained group and 0.34 for the placebo-trained group, a difference of 0.14. Obviously, this information is available only from separate analyses of the glance distributions across glance sequences.

For these analyses, we collapsed the data across the different tasks in order to increase the number of data points used to generate the CDF. However, an examination of the various tasks at different glance thresholds could offer additional insight, as suggested by Figure 2. As shown, the proportion of especially long glances remains relatively high for the map task throughout the entire sequence, whereas the proportion of especially long glances rises and then falls for the other tasks.

In order to further explore task differences, we carried out a task analysis for the various tasks in Table 1. For each task, the subtasks were identified and then coded in terms of the visual demand (after Wickens, 2002; Horrey & Wickens, 2003; see Table 1 for details). The task analysis and coding exercise support the notion that visual demands are higher when information access effort becomes elevated (e.g., when eye and/or head movements are required) and when visual information is more abundant or more difficult to discern (e.g., visual clutter or poor legibility).
More specifically, they also support the notion that visual demands are much higher in the map task throughout than for the other tasks, corroborating the pattern shown in Figure 2. Although the mapping of the subtasks to the glances in the sequence can only be approximated here, we argue that such an approach can effectively complement the sequential analysis of glance behavior by providing more resolution concerning the momentary task demands.

With respect to the illustrative data set employed here, the shift of the distributions in Figure 1 to the right for the later glances suggests one of two plausible, though not mutually exclusive, explanations. First, drivers in both the trained and untrained group make earlier glances for orienting but later glances for resolving and processing relevant information for a task—an outcome that is also supported by the task analysis. Second, drivers’ subjective estimates of the passage of time (their “internal clock”) when they are glancing down may vary as a function of the sequence of glances, being relatively accurate early in the sequence and less accurate (more capable of distraction) later in the most demanding part of the sequence. This variation might occur because drivers are diverting more resources to the in-vehicle task and fewer resources toward an internal clock at each successive glance. When the visual demands decrease later in the sequence for all but the map task, the frequency of especially long glances also decreases, either because the information that is obtained per glance is itself smaller or, as a corollary, because the participants can now devote more resources to the internal clock. Although we cannot resolve these possible explanations here, we underscore that it is the analysis of the distribution of the duration of separate glances in a sequence and the comparison across tasks that can shed some important light on the matter.

In conclusion, analyzing the distribution of glance durations for each glance in a sequence across tasks can yield important information regarding the effects of different interventions (e.g., training). Additionally, comparing the patterns of the sequence distributions of glance durations across tasks can yield important information regarding the effect of the different tasks on the glance distributions at various points in the sequence, possibly leading to changes in the design of an interface to an in-vehicle task. Of special concern for the practitioner are glances in a sequence that occur following the initial glances (e.g., Lee et al., 2012). Sometimes these will be the glances latest in the sequence; oftentimes they will be the glances midway in the sequence, as our task analysis suggests.

**ACKNOWLEDGMENTS**

Thanks are due to John Lee and three anonymous reviewers for helpful comments on an earlier draft of the manuscript. This research is supported by grants: NSF SBR-9413733 (Major Research Instrumentation Grant), DOT TIER 1 UTC (Crash Imminent Safety), and NSF 1405550 RI: Medium (Semi-Autonomous Vehicles) to Donald L. Fisher (as PI, senior investigator, and co-PI, respectively) and Liberty Mutual-UMass Amherst Post-Doctoral Research Fellowship awarded to Yusuke Yamani.
The analysis of the distribution of sequential glance sequences relative to untrained drivers. The results indicated that long glances later in a sequence are almost twice as likely as long glances early in a sequence. These results indicate that trained drivers maintain their improvement in both early- and late-sequence glances relative to untrained drivers. The analysis of the distribution of sequential glance durations, coupled with an understanding of momentary task demands, can shed some important light on in-vehicle glance behavior.

REFERENCES

Yusuke Yamani is an assistant professor in the Department of Psychology at Old Dominion University. He earned his PhD in psychology at the University of Illinois at Urbana-Champaign in 2013.

William J. Horrey is a senior research scientist in the Center for Behavioral Sciences at the Liberty Mutual Research Institute for Safety in Hopkinton, Massachusetts. He received his PhD in engineering psychology from the University of Illinois at Urbana-Champaign in 2005.

Yulan Liang is a research scientist in the Center for Behavioral Sciences at the Liberty Mutual Research Institute for Safety in Hopkinton, Massachusetts. She earned her PhD in industrial engineering at the University of Iowa in 2009.

Donald L. Fisher is a professor in and head of the Department of Mechanical and Industrial Engineering at the University of Massachusetts, Amherst. He received his PhD in psychology from the University of Michigan in Ann Arbor in 1982.

Date received: December 16, 2013
Date accepted: October 18, 2014