FACTORS AFFECTING GLANCE BEHAVIOR WHEN INTERACTING WITH IN-VEHICLE DEVICES: IMPLICATIONS FROM A SIMULATOR STUDY

Yiyun Peng[1], Linda Ng Boyle[1], Mahtab Ghazizadeh[2] & John D. Lee[2]
[1]University of Washington, Seattle, Washington, USA
[2]University of Wisconsin—Madison, Madison, Wisconsin, USA
Email: yiyunp@uw.edu

Summary: This study examined the effects of text entry and reading on drivers' eye glance behavior, as influenced by text length and presence of ambient text (i.e., text around targeted text). A simulator study was conducted with 28 drivers. The findings showed that text entry tasks required longer eyes-off-road (EOR) time than text reading tasks. The presence of ambient text also increased the total EOR time for text reading. Tasks with shorter text required shorter individual glances, but even the shortest text entry tasks resulted in long glances for those who entered text in large chunks. Thus, shortening the text length alone may not ensure safe glance behavior and other countermeasures may need to be considered.

INTRODUCTION

In-vehicle visual-manual distraction undermines driving safety because it diverts drivers' visual attention from the roadway and increases the chance that drivers miss critical events such as lane drifting, lead vehicle braking, or pedestrian crossing. Off-road glances exceeding two seconds during a safety-critical event doubles the risk of crashes/near-crashes (Klauer et al., 2006). Controlled experiments also showed that long off-road glances lead to larger lane deviations and slower responses to lead-vehicle braking (Dingus et al., 1989; Donmez et al., 2007). In general, individual glances that exceed two seconds have been considered as critical for secondary tasks. Therefore, designing in-vehicle displays that reduce drivers' off-road glances is essential for car manufactures to improve driver safety.

Numerous studies have compared the effects of different in-vehicle secondary tasks on drivers' glance behavior (e.g., navigation systems versus radio tuning). However, few studies have examined the various design specifications associated with the same in-vehicle tasks. One such specification is the number of characters to be read or entered for text reading or text entry. Another important aspect relates to the ambient text, or the irrelevant text around the targeted text. Too much ambient text can clutter the display and impact the speed and accuracy of the visual search for the targeted text (Wolfe & Pashler, 1988). Further, when the relevant information (or targeted text) appears consistently in a given information channel or area of interest, observers will tend to sample this channel more often (Horrey et al., 2006). Therefore, if the location of the irrelevant text is randomly placed, it is more likely to draw even more attention away from the driver and pose a greater distraction. The objective of the current study is to examine the effects of ambient text and text length on drivers' eyes-off-road (EOR) for text entry and reading tasks in a simulator study, which can guide designs of in-vehicle systems.

METHODS

Participants

The participants in this study were a representative sample of the U.S. driving population from four age groups (18-24, 25-39, 40-54, and 55-75 years old). Twenty-eight volunteers (seven in each age group, 15 males and 13 females) with a valid drivers' license were recruited from Seattle, WA area. They were native English speakers, in good general health conditions, and reported driving a minimum of 7000 miles per year. They indicated that they were comfortable using computers, touch-screens, and communicating via text messages.

Experimental Design

The effects of task characteristics were examined using a 2 (task type) × 3 (task length) × 2 (ambient text) within-subject design (12 test conditions). The task type had two levels: text entry and text reading. The text entry tasks required participants to enter a word using the touchscreen keyboard and the text reading tasks asked them to read non-scrolling phrases (Table 1). Task length had three levels: short, medium, and long. The three text entry lengths were 4, 6, and 12 characters, and the three text reading lengths were 20 to 40, 60 to 80, and 120 to 140 characters, respectively. Ambient text had two levels (Figure 1): ambient text present (Yes) or ambient text absent (No). Each participant completed three trials of each test condition in two drives.

Table 1. Protocol for the text entry and text reading tasks

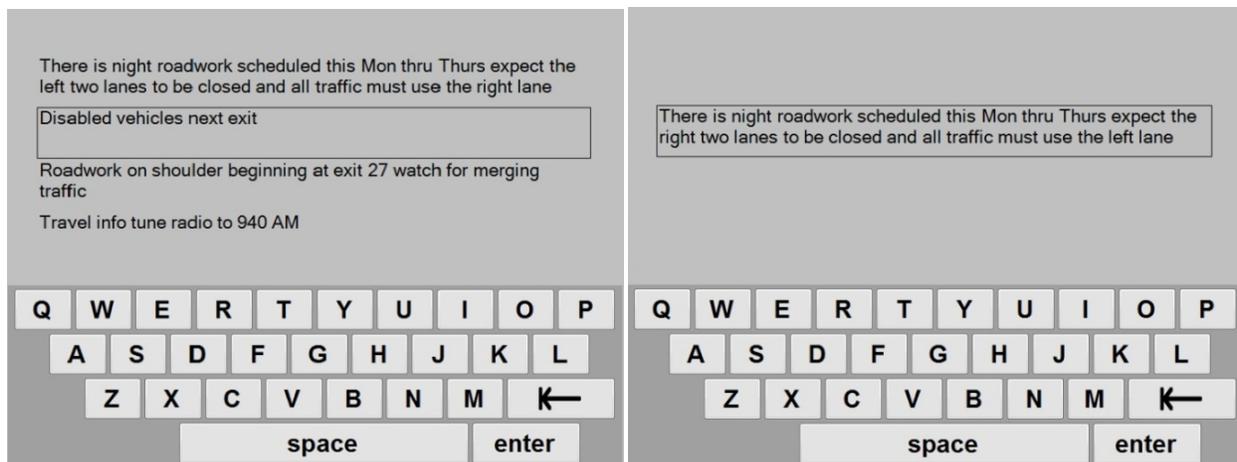| Text Entry	Text Reading
• The computer says a word • User enters keystrokes for word • User presses ENTER when complete	• The computer displays a phrase (in box) • User presses ENTER after comprehending • The computer reads a statement • User presses True or False

Figure 1. Example of screens with (left) and without (right) ambient text

Apparatus

The experiment was conducted using the NADS MiniSim, a low-cost, low-fidelity driving simulator and a Seeing Machines faceLAB eye tracker. The NADS MiniSim included three screens (3.0' [wide] by 1.7' [tall] each) that were placed about 4.5' away from driver's eye point. A 7" touchscreen display with QWERTY keyboard included only keys needed for the tasks (capital alpha characters, backspace, enter, and space bar). There was no type-ahead feature built into the keyboard. The touchscreen display was attached on the right side of the simulator dashboard. All data were collected at 60Hz. The simulated road was a four-lane undivided rural road with a speed limit of 55 mph and no horizontal or vertical curve. All tests were performed while driving in the left lane with no lane changes. Drivers were asked to maintain a speed of approximately 50 mph and a 2-second headway while following a lead vehicle travelling between 52.25 and 55.75 mph. No lead vehicle braking events were included in the scenario.

Procedures

Participants underwent training and a practice drive after consent. In each simulator drive, participants performed 18 tasks (9 readings and 9 entries) in a random order, and each task was 15 seconds apart. Each drive lasted approximately 15 to 20 minutes including the baseline driving. Participants also completed driver demographic, simulator realism, and wellness surveys.

Data Analysis

Repeated measures ANOVAs (using R2.12.1, package 'nlme') were conducted to examine the effect of text entry and reading on the maximum glance duration (Max-GD) and total EOR time (TEORT). Max-GD was defined as the longest duration of all individual glances, and TEORT was the sum of all individual glance durations within the duration of a task (i.e., from the moment the task started to the moment the participant pressed ENTER). Both response variables were log transformed to meet the linear regression assumptions. The final models were chosen based on the lowest AIC values. The majority of drivers did not reach the target driving speed and following distance in the first task of each drive, thus this first task was removed from the subsequent analysis. Two additional tasks were also removed due to unexpected lane changes.

Independent variables included the within-subject factors task type, task length, and ambient text. Preliminary data analysis revealed large individual differences for glance behavior. Therefore, a hierarchical cluster analysis was conducted on the maximum glance duration, average number of glances, and average proportion of EOR, for each driver regardless of test condition, to identify groups of drivers. Euclidean distance and Ward's method was used for the cluster analysis. The driver groups were included in the linear mixed models to account for individual differences.

RESULTS

The mean age of participants in the four age groups were 21.3 ($SD = 0.5$), 32.4 ($SD = 5.4$), 46.8 ($SD = 4.6$), and 59.7 ($SD = 3.0$) years old, respectively. All participants had at least some college education and no crashes in the past year. The cluster analysis revealed two groups of drivers. A group of eight drivers was labeled as high-risk because they had significantly longer maximum

glance duration, smaller average number of glances, and larger percentage of EOR (Figure 2). There were an equal number of females and males in the high-risk driver group. With respect to age, three drivers were between age 18 to 24, two drivers were between age 25 to 39, two drivers were between age 40 to 54, and 1 driver was more than 55 years old. The remaining 20 drivers were labeled as low-risk drivers, and made more frequent but shorter glances to the road and spent a greater proportion of time looking toward the road.
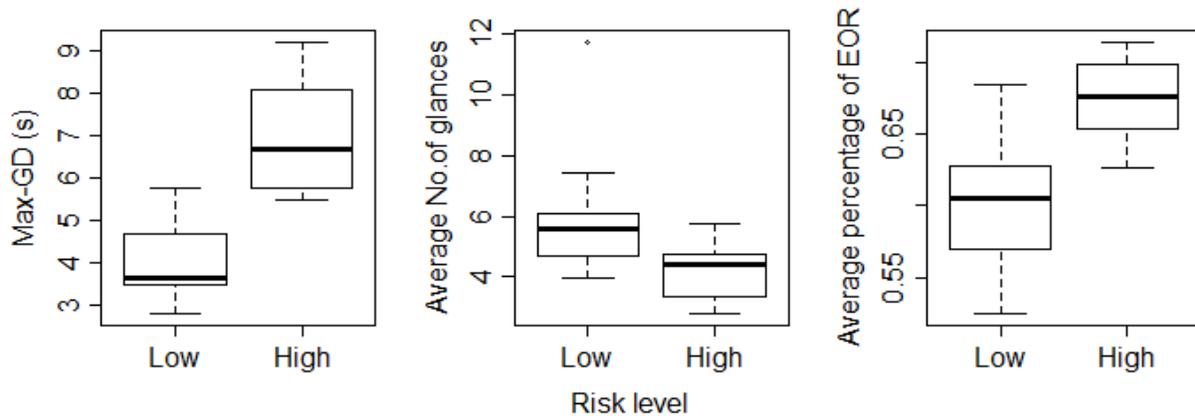

Figure 2. Low and high-risk driver groups based on eyeglance behavior

Text entry took longer to complete ($M = 15.8$ s, $SD = 8.0$ s) than text reading ($M = 7.6$ s, $SD = 5.2$ s). Long text lengths took longer to complete than short and medium text lengths for both text entry ($M = 11.7$ s, 14.4 s, and 23.5 s for short, medium, and long length) and text reading tasks ($M = 5.2$ s, 8.0 s, and 10.7 s, for short, medium, and long length, respectively).

Maximum Glance Duration (Max-GD)

Task type ($F(1, 298) = 308.25$, $p < .0001$), text length ($F(1, 298) = 64.12$, $p < .0001$), and driver groups ($F(1, 26) = 28.41$, $p < .0001$) strongly affected Max-GD. The two-way interaction between task type and text length ($F(2, 298) = 5.53$, $p = .004$), and task type and driver groups ($F(1, 298) = 38.56$, $p < .0001$) were also significant. Additionally, there was a significant three-way interaction between task type, test length, and driver groups ($F(2, 298) = 6.74$, $p = .001$). The two-way interaction between text length and driver group was not significant ($F(2, 298) = 2.80$, $p = .06$). Ambient text did not enter in the final model. Specifically, the model shows that the Max-GD was much longer for text entry than text reading at comparable text length levels, and this increase was even larger for high-risk drivers (Figure 3). When the text reading phrase was short (20 to 40 characters), the estimated maximum glance duration was shorter than 2 seconds for both high- and low-risk drivers. From short to medium and long text reading tasks, the maximum glance duration increased 0.9 and 1.2 seconds for high-risk drivers, but only 0.4 and 0.5 seconds for low-risk drivers. There is no significant difference on Max-GD between medium and long text reading for both groups of drivers (Table 2). For text entry tasks, the maximum glance duration was on average 1.3 seconds longer for high-risk drivers than low-risk drivers. On average, the maximum glance duration increased 0.6 seconds from short (4 characters) to long (12 character) text entry tasks. However, this increase was only statistically significant for the low-risk driver group.
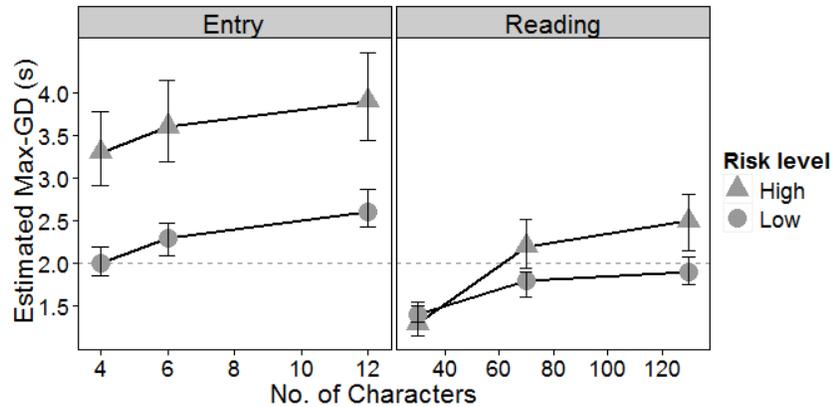
Figure 3. Estimated effects on maximum glance duration (transformed back to the original scale; the error bar shows 95% CI)

Table 2. Tukey contrasts for log(Max-GD)

Driver sub-group	Task Type	Text Length	Estimate	s.e.	z value	Pr(>|z|)
High Risk	Entry	Med vs. Short	0.09	0.07	1.26	0.981
		Long vs. Short	0.17	0.07	2.28	0.462
		Long vs. Med	0.07	0.07	1.03	0.996
	Reading	Med vs. Short	0.52	0.07	6.98	<0.01
		Long vs. Short	0.63	0.07	8.42	<0.01
		Long vs. Med	0.11	0.07	1.45	0.947
Low Risk	Entry	Med vs. Short	0.12	0.05	2.60	0.255
		Long vs. Short	0.27	0.05	5.81	<0.01
		Long vs. Med	0.15	0.05	3.19	0.056
	Reading	Med vs. Short	0.21	0.05	4.45	<0.01
		Long vs. Short	0.29	0.05	6.38	<0.01
		Long vs. Med	0.09	0.05	1.86	0.764

Total Eyes-Off-Road Time (TEORT)

The three main effects, task type ($F(1, 301) = 770.77$, $p < .0001$), task length ($F(2, 301) = 452.21$, $p < .0001$) and ambient text ($F(1, 301) = 20.09$, $p < .0001$) were significant. There was also a significant two-way interaction between task type and length ($F(2, 301) = 8.46$, $p = .0003$). The TEORT was not significantly different between high- and low-risk drivers.
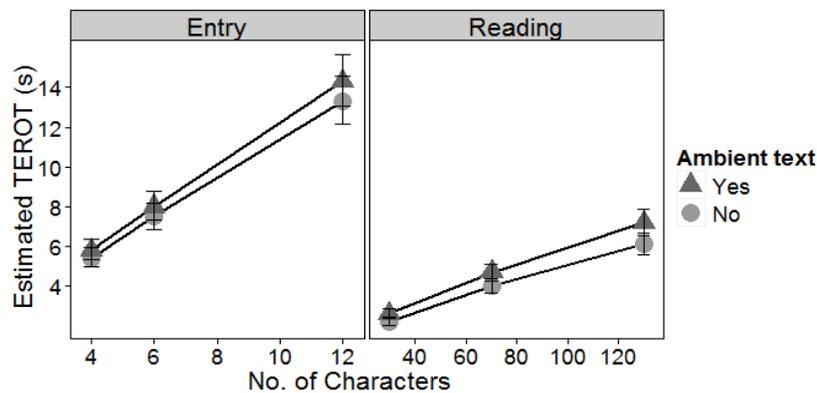

Figure 4. Estimated TEORT (transformed back to the original scale; the error bar shows 95% CI)

For text entry, there was a larger increase in TEORT from medium to long (d = 6.3 s, 79% increase) than from short to medium (d = 2.2 s, 38% increase) tasks (Figure 4). For text reading, the TEORT is expected to increase about 80% (d = 2.1 s) from short to medium length, and 53% (d = 2.5 s) from medium to long length. However, these increases appear to be linear when mapping the length category to the actual number of characters for text entry and reading tasks. TEORT was 0.4 to 1.1 seconds (7.1% increase) longer in tasks with ambient text compared to those without ambient text, and this effect size was estimated to be larger for text reading than text entry but was not significant (F(1, 301) = 3.04, p = .08).

DISCUSSION

Crash risk might not be best represented by the mean of a distribution, but by the extreme glances (Horrey & Wickens, 2007). Therefore, it is important to examine the maximum glance durations that occur during an in-vehicle task to better assess its risk. Large individual differences in glance durations may reflect drivers' risk taking propensity (Donmez et al., 2010), and the chunking techniques that drivers employ. To account for these large individual differences, a cluster analysis identified a group of drivers that appeared to exhibit high-risk behavior in terms of off-road glances. Although this was a small group (8 drivers), their higher risk propensity and longer glances toward in-vehicle displays suggest that they could pose risks for themselves as well as other road users. With respect to the low-risk drivers, the findings suggest that this group can maintain 2 seconds or shorter individual glances while reading text messages that are as long as 140 characters, or entering text that are 4 characters long. For these same tasks, the maximum individual glances are much longer for the high-risk drivers. For high-risk drivers, shortening the text entry tasks does not guarantee safe EOR durations and additional countermeasures may be needed (Lee, 2009). For example, given that the long individual glances were likely to be caused by drivers typing several letters at once, properly designed type-ahead feature and real-time alerting systems may help reduce the length off-road glances and bring drivers' eyes back to the road more often.

The presence of ambient text slightly increased the TEORT, but not the maximum individual glance durations. Drivers need to alternate their glances between the road and the touchscreen while driving, and each time, they must relocate the text box on the screen to continue reading/entering text. When ambient text is present, it may take longer to separate the text box from the clutter of the surrounding text, especially for the reading tasks, where the overall text density is greater than the text entry tasks. It should be noted that there are many possible layouts for ambient text, and the current study only examined one possible layout. Understanding the effect of other types of ambient text warrants further investigation.

Given that this was a simulator study, drivers may have driven in a way that is not representative of how they might drive on the road. The effects of distraction also depend on experimental protocols and outcome measures. For example, the low demands of simulated driving tasks could be one reason for the extremely long off-road glances (e.g., greater than 8s), which may not represent driving in actual road conditions. In this study, researchers reminded participant not to lag too far behind the lead vehicle, but there was no automatic feedback regarding the appropriate headway distance as was provided by Ranney et al (2012). Drivers' eye glance behavior might also depend on the complexity of the driving environments (e.g., opposing traffic,

stop and go traffic) and this should be examined in future research to examine the magnitude of risk taking among drivers.

Selection of outcome variables and data aggregation methods can also affect findings. The drivers with extreme glance behaviors did not have significantly longer TEORT, which may indicate that individual glance duration is better for assessing risk taking. Studies that examine in-vehicle secondary tasks should be able to identify behaviors of the few risky drivers, and design countermeasures to reduce distraction and thus improve roadway safety for all users.

ACKNOWLEDGMENTS

REFERENCES

Dingus, T., Antin, J., Hulse, M., & Wierwille, W. (1989). Attentional demand requirements of an automobile moving-map navigation system. Transportation Research Part A, 23(4), 301-315

Donmez, B., Boyle, L., & Lee, J. (2007). Safety implications of providing real-time feedback to distracted drivers. Accident Analysis & Prevention, 39(3), 581-590

Donmez, B., Boyle, L., & Lee, J. (2010). Differences in off-road glances: effects on young drivers' performance. Journal of Transportation Engineering, 136(5), 403-409

Horrey, W., & Wickens, C. (2007). In-vehicle glance duration: distributions, tails, and model of crash risk. Transportation Research Record, 2018(1), 22-28.10.3141/2018-04

Horrey, W. J., Wickens, C. D., & Consalus, K. P. (2006). Modeling drivers' visual attention allocation while interacting with in-vehicle technologies. Journal of Experimental Psychology: Applied, 12(2), 67

Klauer, S., Dingus, T., Neale, V., Sudweeks, J., & Ramsay, D. (2006). The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic driving study data. Washington, DC.

Lee, J. (2009). Can technology get your eyes back on the road? Science, 234, 344-346

Ranney, T., Baldwin, S., Smith, L. A., Martin, J., & Mazzae, E. (2012). Driver behavior during visual-manual secondary task performance: occlusion method versus simulated driving. Washington, DC: National Highway Traffic Safety Administration.

Wolfe, J. M., & Pashler, H. (1988). Visual search. In Attention. London: University College London Press.