Differences in Off-Road Glances: Effects on Young Drivers’ Performance

Birsen Donmez¹; Linda Ng Boyle²; and John D. Lee³

Abstract: Young drivers display more risk-taking behavior than other age groups. Performing distracting tasks is one such risky behavior that is observed among young drivers. However, due to inexperience, young drivers may not be able to appropriately compensate for the effects of distractions. A driving simulator study with 53 young drivers (aged 18–21) was conducted to assess the level of engagement with an in-vehicle secondary task. A cluster analysis revealed three groups of drivers that differed based on eye glance behavior and driving performance: drivers with low-risk, moderate-risk, and high-risk behavior. A subset of these drivers was provided with feedback to help modulate their distracting activities with the riskiest group benefitting most from feedback as indicated by enhanced glance behavior and driving performance. The findings have implications for developing better crash countermeasures to mitigate the effects of distraction.

DOI: 10.1061/(ASCE)TE.1943-5436.0000068

CE Database subject headings: Accident prevention; Technology; Traffic safety; Human factors; Driver behavior.

Author keywords: Accident prevention; Technology; Safety; Human factors; Driver behavior.

Introduction

Driving involves complex interactions between the driver, vehicle, and environment. Breakdowns in any of these interactions undermine driving safety. The introduction of in-vehicle and carried-in devices (e.g., cell phones and MP3 players) raises concerns that the demands of such systems may conflict with the demands of driving. Driver distraction can be defined as diminished attention of the driver to the driving task. A driver’s willingness to engage in a non-driving task and the attentional demands placed on the driver by that task can contribute to distraction. Vehicular crashes caused by driver inattention and resulting from driver distraction are a major concern (Brookhuis et al. 2003; Regan et al. 2008; Stutts and Hunter 2003; Wang et al. 1996). The growing number of potentially distracting devices can further undermine safety (Verwey et al. 1996) and pose even greater hazards to young drivers (Lee 2007).

Distracting activities have been shown to be a particular risk to young drivers (McGehee et al. 2007; Neyens and Boyle 2007). Younger drivers’ greater sensitivity to distraction may be due to their limited ability to effectively direct their attention to the road-way (Fisher et al. 2002). Inexperienced drivers make longer fixations and scan smaller areas of the visual scene compared to experienced drivers (Mourant and Rockwell 1972). With experience, their scanning adjusts to reflect the spatial-temporal characteristics of hazardous situations (Brown and Groeger 1988). As an example, drivers’ scanning becomes more sensitive to road type with experience (Underwood et al. 2003). Both crash data as well as on-road and simulator studies show that crashes among young drivers are a result of failures of attention and visual search (McKnight and McKnight 2003; Neale et al. 2005). With respect to sharing the attentional demands of the roadway with a secondary task, Wikman et al. (1998) showed that young and inexperienced drivers tend to look away from the road with longer and more variable glances compared to those with more experience. Specifically, 29% of young drivers had glances longer than three seconds, while the more experienced drivers had no glances as long.

Young drivers have a higher risk for crash involvement, in part because they are generally more likely to take risks while driving (Deery 1999; Ferguson 2003; Williams 2003). They may also be particularly vulnerable to distractions because of their greater propensity to engage in distracting activities (Olsen et al. 2005). In fact, younger drivers are more likely to use their cell phones while driving than other age groups (Glassbrenner 2005). Redelmeier and Tibshirani (1997) estimated that young drivers are 6.5 times more likely to be involved in a crash when using a cell phone compared to when not using a cell phone. Substantial evidence suggests that stable attitudes and behavioral differences influence crash involvement (Parker et al. 1992). Specifically, a survey of driving behavior among college students showed that those students who reported a high level of sensation seeking were more likely not to wear seat belts and to drive aggressively. These drivers were also more likely to report that they would drive faster on highways and on wet roads when driving a vehicle with antilock brakes (Jonah et al. 2001). A survey of 198 drivers between the ages of 16 and 19 revealed several distinct types of drivers, which varied according to risk-taking propensity. Such
risk-taking included speeding, driving-related aggression and hostility (Deery and Fildes 1999). A subsequent simulator study showed impaired attention-management performance of these drivers in high-workload situations (Deery and Fildes 1999). Overall, individual differences associated with risk-taking may have a strong effect on the degree to which young drivers engage in distracting activities.

Appropriate feedback can help diminish both the impact and the amount of risk-taking behavior (Donmez et al. 2007, 2008a,b). The advances in technology make it possible to provide feedback based on the driver as well as roadway state. For example, eye tracking systems allow us to alert the drivers if their eye glances toward a secondary task exceed a particular threshold, and vehicle sensors enable feedback to be designed based on how far a driver deviates from the center of their primary lane. Ideally, a system that provides feedback concerning risky glance behavior would help drivers moderate the degree of distraction they are willing to engage. The objective of this study is to identify risky young drivers based on their glance behavior, and to assess how well feedback guides risky drivers to more appropriate glance behavior and improved driving performance.

Method

Participants
There were 53 participants in this study between the ages of 18 and 21 (female: n=26, \(\bar{X}=19.6, sd=1.12\); male: n=27, \(\bar{X}=19.6, sd=1.08\)). Participants were paid $15 per hour, and then had the opportunity to earn a bonus of up to $15. Drivers were recruited through newspaper advertisements and had to be frequent drivers (i.e., they drove more than three days a week) with at least two years of driving experience.

Apparatus
The experiments were conducted with a medium-fidelity, fixed-based simulator with a 50° visual field powered by Global Sim, Inc.’s DriveSafety Research Simulator. All graphics for roadway layouts, markings, and signage conform to AASHTO and Manual of Uniform Traffic Control Devices design standards. The cab is equipped with forward feedback steering wheel, actual gauges, and a rich audio environment to enhance realism. A 7-in. LCD (60-Hz frame rate at 640×480 resolution) mounted on the dashboard by a small stand was used in the experiment for the presentation of the visual messages used in the secondary task. The viewing angle from the driver’s eye point was approximately 18°. A Seeing Machines eye tracker was used to collect eye movement using FaceLab 4.2: an eye and head tracking system that enables analysis of natural behavior by using a set of cameras as a passive measuring device.

Experimental Design
Each participant completed one identical experimental drive (approximately 7 min long). The drive took place on a two-lane rural road with oncoming traffic. Participants were instructed to drive at 72 km/h (45 mph), and to follow a lead vehicle that periodically braked mildly (0.2 g) for 5 s. The lead vehicle braking events were considered moderately demanding, whereas the rest of the driving scenario was low demand. For experimental control, the lead vehicle speed was adjusted to maintain a 1.8 s head-way time before a lead vehicle braking event, otherwise the lead vehicle speed was set to 72 km/h. There were a total of 10 braking events in the drive. Driving performance was assessed by averaging the minimum time-to-collision (TTC) values over these 10 events. Minimum TTC is defined as the minimum time it takes to collide with an object given the instantaneous relative speed and distance. TTC has been proposed and used as a crash-avoidance metric in forward collision avoidance systems (Minderhoud and Bovy 2001; Vogel 2003) and is associated with the likelihood of collision (Lee et al. 2002). Minimum TTC is the shortest TTC during a braking event, and has been widely used to assess driver performance (Belz et al. 2004; Donmez et al. 2006; Young et al. 2008).

Participants completed this drive while performing a self-paced in-vehicle secondary task, which has been shown to undermine driving performance (Donmez et al. 2007). The task was designed to simulate visual, motor, and cognitive distractions typical of many in-vehicle system interactions (e.g., scanning an MP3 play-list). Participants were instructed to interact with the secondary task whenever they felt comfortable, in order to ensure that eye movement patterns were an indicator of willingness to engage in the distraction. There was also monetary compensation based on speed and accuracy to provide an incentive to engage in this task. The monetary compensation was provided to increase the secondary task importance for the participants to simulate real-life interactions with in-vehicle systems (e.g., a phone conversation with one’s boss). For these reasons, all drivers engaged in the distracting task, but at varying levels.

After participants completed the necessary Internal Review Board consent forms, they were asked to complete one practice drive to become familiar with the simulator controls and the secondary task. All participants then completed the experimental drive. The data from this single drive is used to cluster the drivers based on their eye movement patterns (hence how they engage in distractions). A glance to the in-vehicle display was recorded when the driver’s gaze vector intersected with the in-vehicle display.

Cluster Analysis
Ward’s Hierarchical Clustering method was used to find groupings in the data. This analytical technique partitions a data set into different groupings (clusters), on the basis of similarities and distances (Johnson and Wichern 2002). Cluster analysis has previously been used to examine transportation issues related to transport risk perception (Olteadal and Rundmo 2007), travel patterns (Ma and Goulias 1997), drivers’ response to collision warnings (Lee et al. 2002), and drivers’ propensity to change behavior based on driver information systems (Conquest et al. 1993).

Clustering of data was performed using the PROC CLUSTER procedure in SAS 9.1. Distances between clusters were computed using Ward’s minimum-variance method, an approach based on the minimization of within-cluster distances. Since the variables with large variances tend to have more effect on the resulting clusters than those with small variances, the data was standardized with the inclusion of STD option in PROC CLUSTER. Ward’s Hierarchical Clustering is an agglomerative (bottom-up) clustering technique. As such, there are no limits imposed on the cluster generation. Rather, the number of clusters is identified by the semipartial R-squared values at different cluster levels. This value measures loss of homogeneity due to merging and will be small when two clusters are very similar. The number of clusters
receiving any feedback. The two variables used to cluster the drivers were the number and the mean duration of glances to the in-vehicle display in this single drive. These variables produced three clusters [Fig. 1(a)] that accounted for over half of the variation in the data (60%). Statistical analysis on eye movement patterns revealed that Cluster A \((n=20)\) had the longest mean duration of glances to the in-vehicle display when compared to Clusters B \((n=15)\) and C \((n=18)\) [Table 1, Fig. 1(b)]. The glance frequencies were the largest for Cluster B followed by Cluster A and then Cluster C.

In terms of labeling these clusters, it would appear that drivers in Cluster A tend to be riskier with longer off-road glances compared to Clusters B and C. Cluster B had higher number of off-road glances, but shorter duration off-road glances than Clusters A and C. Hence, Cluster C appears to be the least risky of the three.

Performance tends to decline if drivers have many long in-vehicle display glances. However, many short in-vehicle display glances do not necessarily degrade performance. Horrey and Wickens (2007) indicate that unsafe conditions leading to a crash reside not at the mean of a distribution but rather in the tails. Therefore, to confirm the initial labeling of the clusters, particularly long in-vehicle glances were also considered. The 95 percentile glance durations to the in-vehicle display over the drive was examined with respect to these cluster groups (Table 1, Fig. 2). The results align with those obtained for the mean glance duration. That is, Cluster A had the longest 95 percentile glance durations when compared to Clusters B and C.

Cluster C had significantly longer minimum TTC during lead vehicle braking events when compared to Cluster A (Table 1, Fig. 3). The estimated value for Cluster B was between these two clusters but was not significantly different from either. Cluster C reflects safer driving behavior with relatively few and short off-road glances. Henceforth, the clusters are labeled as low-risk (Cluster C), moderate-risk (Cluster B), and high-risk drivers (Cluster A).

### Table 1. Statistical Comparisons among the Three Clusters

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Estimate ((\Delta))</th>
<th>(t)-value</th>
<th>(df)</th>
<th>(p)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean glance duration to the in-vehicle display</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster A (high risk) versus C (low risk)</td>
<td>0.34s</td>
<td>7.63</td>
<td>50</td>
<td>&lt;0.0001</td>
<td>(0.25, 0.43)</td>
</tr>
<tr>
<td>Cluster A (high risk) versus B (moderate risk)</td>
<td>0.36s</td>
<td>7.71</td>
<td>50</td>
<td>&lt;0.0001</td>
<td>(0.27, 0.46)</td>
</tr>
<tr>
<td>Cluster B (moderate risk) versus C (low risk)</td>
<td>-0.02s</td>
<td>-0.45</td>
<td>50</td>
<td>0.66</td>
<td>(-0.12, 0.07)</td>
</tr>
<tr>
<td><strong>Number of glances to the in-vehicle display per minute</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster A (high risk) versus C (low risk)</td>
<td>3.36</td>
<td>2.81</td>
<td>50</td>
<td>0.007</td>
<td>(0.96, 5.77)</td>
</tr>
<tr>
<td>Cluster A (high risk) versus B (moderate risk)</td>
<td>-6.78</td>
<td>-5.38</td>
<td>50</td>
<td>&lt;0.0001</td>
<td>(-9.31, -4.25)</td>
</tr>
<tr>
<td>Cluster B (moderate risk) versus C (low risk)</td>
<td>10.14</td>
<td>7.87</td>
<td>50</td>
<td>&lt;0.0001</td>
<td>(7.56, 12.73)</td>
</tr>
<tr>
<td><strong>95 percentile glance duration to the in-vehicle display</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster A (high risk) versus C (low risk)</td>
<td>0.79s</td>
<td>8.09</td>
<td>50</td>
<td>&lt;0.0001</td>
<td>(0.59, 0.97)</td>
</tr>
<tr>
<td>Cluster A (high risk) versus B (moderate risk)</td>
<td>0.84s</td>
<td>8.22</td>
<td>50</td>
<td>&lt;0.0001</td>
<td>(0.64, 1.05)</td>
</tr>
<tr>
<td>Cluster B (moderate risk) versus C (low risk)</td>
<td>-0.05s</td>
<td>-0.51</td>
<td>50</td>
<td>0.61</td>
<td>(-0.26, 0.16)</td>
</tr>
<tr>
<td><strong>Minimum TTC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cluster A (high risk) versus C (low risk)</td>
<td>-2.16s</td>
<td>-2.23</td>
<td>50</td>
<td>0.03</td>
<td>(-4.12, -0.21)</td>
</tr>
<tr>
<td>Cluster A (high risk) versus B (moderate risk)</td>
<td>-0.82s</td>
<td>-0.80</td>
<td>50</td>
<td>0.43</td>
<td>(-2.87, 1.23)</td>
</tr>
<tr>
<td>Cluster B (moderate risk) versus C (low risk)</td>
<td>-1.35s</td>
<td>-1.29</td>
<td>50</td>
<td>0.20</td>
<td>(-3.45, 0.76)</td>
</tr>
</tbody>
</table>

---

**Fig. 1.** (a) Representation of each driver in the glance space based on cluster membership; (b) the distribution of eye movement measures toward in-vehicle display by cluster group.
Effect of Feedback

The type of feedback tested was a combined concurrent (provided in real-time) and retrospective feedback (provided after a trip), which alerted the drivers of high distraction levels and of incidents that may have occurred (e.g., lane deviations). A previous study showed that combined concurrent and retrospective feedback can enhance driving performance and behavior (Donmez et al. 2008b).

To examine the effect of feedback, 27 of the 53 drivers that completed the first drive were observed while driving the same route seven more times. They performed these drives while engaging in the same secondary task, with participants randomly assigned to receive feedback or not. Drivers in the feedback condition, there were six drivers from the high-risk group, three drivers from the moderate-risk group, and six drivers from the low-risk group. In the feedback condition, there were six, five, and four drivers in high-, moderate-, and low-risk groups, respectively.

The analysis of mean glance duration to the in-vehicle display revealed that feedback altered glance behavior of the high-risk group (Fig. 4). In particular, high-risk drivers who received feedback had shorter mean glance duration to the in-vehicle display when compared to high-risk drivers who did not receive feedback ($F(1,20.5)=10.79, p=0.004$). For this group of drivers, feedback decreased mean glance duration by an estimated 0.40 s [95% confidence interval (CI): 0.15, 0.66]. There was no interaction with the drive, and hence this effect was fairly constant over time. With each additional drive, the duration of very long glances for the high-risk drivers increased by 0.15 s when no feedback was provided (95% CI: 0.09, 0.22) and by 0.07 s when feedback was provided (95% CI: 0.02, 0.12). Compared to high-risk drivers who did not receive feedback, with each additional drive, high-risk drivers who received feedback had an estimated 0.09 s shorter 95 percentile glance duration to the in-vehicle display (95% CI: 0.01, 0.17). Therefore, feedback decreased the duration of very long glances of high-risk drivers over time. There were no significant effects of feedback on the number of glances to the in-vehicle display.

Feedback also mitigated the effects of risky driving behavior for high-risk drivers, by increasing the minimum TTC values over time (Fig. 5). Among high-risk drivers, for each additional drive,
feedback increased the minimum TTC by an estimated 0.90 s (95% CI: 0.50, 1.29). At the end of the eighth drive, high-risk drivers who received feedback had an estimated 4.19 s (95% CI: 1.10, 7.28) larger minimum TTC when compared to moderate-risk drivers who received feedback. Feedback did not result in statistically significant effects for the low and moderate-risk groups.

Discussion

This study demonstrates that some young drivers exhibit more risky glance patterns than others. The number and the mean duration of glances to the in-vehicle display can divide drivers into three clusters. One cluster of drivers looked away from the road for the longest periods and because they also had the shortest minimum TTC, was labeled “high risk.” Another cluster group had the fewest off-road glances, which were also relatively short in duration. This cluster was labeled as low risk. The third cluster did not significantly differ in terms of driving performance from that of the low-risk group. However, they did have significantly more off-road glances than the low-risk group and was hence labeled a moderate-risk group since frequent off-road glances may be an issue when there are more demanding tasks than were examined in this study. In these high demand situations, there may be a driving performance decrement and it would be worthwhile to explore this in future studies.

Previous studies have revealed differences in risk-taking between males and females (Evans 2004; Williams and Shabanova 2003). The low-risk cluster consisted of a larger proportion of female drivers (72% female), whereas the moderate-risk cluster was mostly males (27% female). The high-risk cluster had approximately equal number of males and females. Although it was expected that the low-risk cluster would have more females, we also expected the high-risk group to have more males (Laapotti and Keskinen 1998). There are also studies showing gender differences in young drivers among personality traits such as sensation seeking (Jonah 1997; Jonah et al. 2001) and as young drivers begin to accumulate driving experience (Williams 2003). A study compared the crash and conviction rates of 28,500 Finnish novice drivers in age brackets of 18–20, 21–30, and 31–50 years old. When examined in terms of the levels of driving tasks (Michon 1985), the results showed young novice drivers, particularly males, had a greater number of strategic and tactical errors, and that females tended to have greater problems at the operational level of driving behavior (Laapotti et al. 2001).

The findings of this study suggest that drivers differ in how they modulate their attention to secondary tasks and that these differences influence their driving performance. The strong differences between the three groups of drivers suggest that a subgroup of young drivers exist that might benefit from feedback regarding their risky glance behavior. In fact, feedback only had a significant influence on the eye glance behavior of the high-risk group. High-risk drivers who received feedback had shorter mean glances to the in-vehicle display when compared to those who did not receive feedback. Moreover, this influence led to a fairly constant improvement over time. In addition to diminishing risky eye glance behavior, feedback also helped mitigate the effects of such behavior for the high-risk drivers, by helping them achieve larger minimum TTC values over time when responding to lead vehicle braking events. At the end of the eighth drive, high-risk drivers who received feedback had significantly larger minimum TTC values than moderate-risk drivers who also received feedback. The enhancement of lead vehicle braking response over time may be attributed in part to better modulation of attention with feedback. These results suggest that feedback may shift the distribution of attention and enhance driver’s response to roadway events. This may then lead to improving overall driving behavior.

In a study that examined the effects of feedback on training teenage drivers, McGehee et al. (2007) defined risky behavior based on the frequency of observed incidents. McGehee et al. (2007) separated 26 drivers into high and low risk (defined as high and low incident frequency in the study). The seven teenage drivers identified as high risk showed an 89% decrease in the number of incidents by the 18th week of the study. Consistent with the current study, the group of teenagers with risky driving behavior benefited most from feedback.

McKnight and McKnight (2003) suggest that young drivers drive in a risky fashion mostly because they are clueless, rather than careless. That is, a great majority of nonfatal young driver crashes are due to errors in attention and visual search, rather than blatantly risky behavior. The brief exposure to feedback in this current experiment is highly unlikely to shift a stable personality trait such as risk-taking tendency. However, given that feedback was able to help high-risk drivers diminish their risky eye glance behavior, it appears that feedback can in fact help clueless young drivers be more aware of and understand the risks of being distracted.

It is recognized that the effect estimates generated from a driving simulator study may not be directly comparable to on-road performance due to factors such as an experimenter being present or the lack of any actual crash risk in a simulated environment. However, based on the findings of simulator validation studies (Godley et al. 2002), the relative (not absolute) benefits of feedback to different risk groups would still be expected. Moreover, a single experiment cannot test for all the possible driving situations and driver populations. Thus, further research is needed to identify if the results obtained in this study would generalize beyond the specific conditions and drivers tested in this study.

The most effective means of delivering feedback to risky drivers depends on the factors that guide such behavior. If the risky behavior stems from knowingly incurring greater risk at the level of tactical driving behavior, then analysis of drivers’ decision making capabilities and better training might be most effective. If risky behavior stems from poor modulation of attention at the level of operational driving behavior, then real-time feedback that directs drivers’ attention back to the road might be most effective. Thus, further research focusing on the strategic, tactical, and operational levels may provide greater insight into the benefits of different feedback types (Lee et al. 2008). Moreover, a critical concern with feedback is how drivers might respond to feedback with long-term use. Although the results of this experiment are very promising, issues of driver acceptance and willingness to
adapt over the long-term will determine whether providing drivers with feedback can mitigate the growing problem of distraction for young drivers.

Conclusions

This study suggests that young drivers are not a homogenous population and respond to distractions quite differently. Drivers who have many long glances to an in-vehicle display (i.e., risky drivers) perform worse than those who have a few short glances. The results also suggest that feedback can provide great benefits to these risky drivers, by decreasing the glance durations to an in-vehicle display and by enhancing response to lead vehicle braking events. Therefore, feedback can diminish risky glance behavior and also mitigate its effects. Building in-vehicle systems with feedback capabilities can help address the safety concerns associated with introducing potentially distracting technology in the cars. The findings, therefore, have practical implications for systems designed to help mitigate the effects of distraction.

Acknowledgments

This research was partly funded by Dwight David Eisenhower Graduate Fellowship, National Highway Institute, U.S. Department of Transportation, and a National Science Foundation Career Award (Grant No. 0643390). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the writers and do not necessarily reflect the views of the DOT or NSF. Data collection for this project was also provided, in part, by Gregor Veeger. A version of this paper was also presented at the 10th Applications of Advanced Technologies in Transportation (AATT) in Athens, Greece.

References


