SAE INTERNATIONAL®

The Influence of Driver's Age on Glance Allocations during Single-Task Driving and Voice vs. Visual-Manual Radio Tuning

2016-01-1445
Published 04/05/2016

Jonathan Dobres, Bryan Reimer, and Bruce Mehler
MIT AgeLab

James Foley and Kazutoshi Ebe
Toyota Technical Center USA, Inc.

Bobbie Seppelt and Linda Angell
Touchstone Evaluations Inc.

CITATION: Dobres, J., Reimer, B., Mehler, B., Foley, J. et al., "The Influence of Driver's Age on Glance Allocations during Single-Task Driving and Voice vs. Visual-Manual Radio Tuning," SAE Technical Paper 2016-01-1445, 2016, doi:10.4271/2016-01-1445.

Abstract

Driving behaviors change over the lifespan, and some of these changes influence how a driver allocates visual attention. The present study examined the allocation of glances during single-task (just driving) and dual-task highway driving (concurrently tuning the radio using either visual-manual or auditory-vocal controls). Results indicate that older drivers maintained significantly longer single glance durations across tasks compared to younger drivers. Compared to just driving, visual-manual radio tuning was associated with longer single glance durations for both age groups. Off-road glances were subcategorized as glances to the instrument cluster and mirrors ("situationally-relevant"), "center stack", and "other". During baseline driving, older drivers spent more time glancing to situationally-relevant targets. During both radio tuning task periods, in both age groups, the majority of glances were made to the center stack (the radio display). However, compared to visual-manual task periods, during the auditory-vocal periods, significantly more glances were made to situationally-relevant targets and fewer glances to the radio display. These results suggest that, while the auditory-vocal interface pulls some resources away from the forward roadway, it produces glance allocation profiles more similar to baseline driving. As with single-task driving, during the auditory-vocal radio task, older drivers made significantly longer glances off-road (173ms longer, on average), than younger drivers. These findings suggest that the assessment of glance behavior during on-road driving should consider that not all glances away from the forward roadway are necessarily "off-road", i.e. diversions from driving-related attention.

Introduction

Modern technologies increasingly feature voice interfaces for user interaction, such as Apple's Siri, Google Voice, Microsoft's Cortana, and Amazon's Alexa. While these major technology companies represent the leading edge of this technology, voice interfaces can also now be found in something as mundane as a cable television remote control. Many voice implementations are found in the automotive space, and consumers have come to expect access to various complex "infotainment" features from within their vehicles. Such systems include a number of OEM in-dash implementations, which offer a mix of visual-manual and/or touchscreen controls, as well as options for voice command. Likewise, Apple's CarPlay and Google's Android Auto connect a user's smartphone to the in-dash screen to present a specialized interface for the smartphone's functions, and emphasize a mixture of auditory-vocal and visual-manual controls to varying degrees.

Voice interfaces are of particular interest in the automotive sector, since they are generally presumed to allow the driver to perform secondary infotainment functions while more easily maintaining visual attention to the driving scene. Their expected benefit is straightforward-if the driver can perform secondary tasks without having to manipulate buttons or look at screens, this should allow for a driver to be oriented towards the road in a manner more similar to single-task driving. This presumably reduces the potential for distraction and increases attentiveness to the roadway in a way expected to promote safe operation.

While this expected benefit is appealing, data in support of it varies, depending upon the way in which voice systems are implemented, and the way in which they are evaluated. In recent years, the implementation of voice interfaces in the vehicle has evolved as its application has broadened. Not all implementations of voice are the same [9], and "voice-based" tasks should not be treated or evaluated as a single, monolithic group. Voice, as a modality, is used in a variety of ways in modern HMIs, resulting in more complex and multimodal sets of demands on driver resources (demands that are not always just auditory or vocal, but may include visual and touch components as well) [17]. Given how quickly this segment of technology for in-vehicle use is evolving, more up-to-date evaluations of the effects of voice-based tasks are needed.

Over the last fifteen years, several different approaches have been taken to examine the effect of voice interactions on driver behavior and/or cognitive distraction. One of the earliest studies in this area showed that, in a simulated driving task using a production-level interface, both younger (21-35) and older (55-70) participants took less time to complete a variety of tasks when using voice controls as opposed to steering wheel-based controls, but that both control schemes took longer than traditional console controls. Peripheral detection degraded more when using manual controls compared to voice controls, and detection was worse in both conditions compared to baseline driving [3,8]. In another PC-simulator study, it was shown that task completion times were markedly higher for voice interfaces compared to their manual equivalents, though glance behavior was not examined specifically [6]. A field study indicated that voice interfaces, in comparison to visual-manual equivalents, appeared to result in reduced mean off-road glance duration and total off-road glance time, and an associated increase in percentage of time spent looking at the forward roadway [4].

In an early examination of the Ford SYNC™ in-vehicle system [18] in a driving simulator, it was found that voice interfaces were quite advantageous compared to manual controls, producing reductions in task completion time, off-road glance time, and standard deviation of lane position. However, these results should be interpreted with some caution. Approximately 7% of trials were excluded from analysis due to difficulty using the interface. Moreover, the comparison tasks across modalities were not always equivalent. For example, the manual text messaging task asked drivers to type in the message, while the vocal equivalent had users select a message from a predefined list. Another on-road study of the SYNC interface [12] showed that, in comparison to handheld operation, contact dialing and track playing tasks using the voice interface interfered less with vehicle control, were completed faster, and received lower self-reported mental demand scores on the NASA-TLX workload rating. Manual secondary tasks were also associated with more frequent and longer duration glances off the forward roadway, while auditory-vocal equivalents had glance behaviors more consistent with those in baseline single-task driving.

Studies utilizing "Wizard-of-Oz" voice interfaces, in which a voice interface is mimicked through experimenter-deployed audio in response to the driver, have shown a mixture of findings. A test track study [15] showed that, when using a voice interface, several metrics of driver behavior deteriorated, though visual behavior could not be directly assessed owing to unreliability in the eye tracking hardware. A follow-up study, also conducted on a test track [14], showed similar degradations across all voice tasks, though performance was better than when drivers were also asked to use a map for navigation.

At least one study has found that while voice-based tasks took longer to complete than their manual equivalent, there were no significant differences in measures of visual attention, including a peripheral detection task [23].However, this study was conducted on a closed course and may not fully reflect real-world driving conditions.

This brief review indicates that voice controls tend to keep the driver's eyes on the forward roadway to a greater degree than manual controls. However, this behavior is still somewhat different from single-task driving under ordinary driving conditions. The majority of these studies employ driving simulator or test track setups, which do not fully replicate the stresses and nuances of on-road driving in the real world. Therefore, evaluation of contemporary voice-based interfaces under on-road driving conditions would be of interest for understanding how this evolving HMI technology is affecting drivers.

National Highway Traffic Safety Administration (NHTSA) guidelines for the assessment of visual-manual distraction during invehicle task performance make a number of assumptions designed to aid replicable testing [11]. The guidelines assume the use of a driving simulator and a relatively simple driving scene, which in turn affects the assumptions governing the calculation of glance behavior metrics. For example, the guidelines state that any glance off of the forward roadway, whether to the in-vehicle device or to situational targets such as rearview or side mirrors, is to be regarded as an off-road glance: "The duration of an individual glance is determined as the time associated with any eye glances away from the forward roadway. Due to the driving scenario [a simulated environment], eye glances to the side of the roadway or to the vehicle's mirrors are expected to be minimal" (p. 24888, Section VI.F.1). The analytical assumption is perhaps that, during the test evaluation of a task in the prescribed simulator scenario, any glance off of the forward roadway, or the vast majority of them, represents a reallocation of visual attention to a non-driving task. However, in a real-world driving context on the road, such an assumption would likely not apply, since in real world driving behavior, glances to mirrors or the instrument cluster often represent the driver's "situationally-relevant" sampling of information needed to assess traffic, and to perform immediate navigation and headway control tasks-even while interleaving the performance of secondary tasks. Drivers' sampling behavior to collect "situationally-relevant" information has been shown to vary as a function of task type. In an on-road study examining visual scanning when drivers were engaged in a visual-manual task compared to single-task driving, drivers tended to sample different environmental elements to achieve awareness of the driving situation between the two task conditions [24]. Under the higher load imposed from the increased visual demand of visual-manual tasks, drivers placed greater emphasis on, and thus sampled more, elements related to vehicle control tasks. These results are consistent with the more generalized finding that when under task load, even load induced from auditory-vocal tasks, drivers often shed what they believe to be non-critical scanning of peripheral regions and narrow or focus their scanning more tightly on those areas immediately surrounding the vehicle [5,7,21].

This effect of task load inducing drivers to focus their gaze on areas specific to performance of control functions is expected to vary dependent on expertise. As compared to novice drivers, more experienced drivers are known to concentrate their gaze in a smaller area on the forward roadway and to increase their sampling frequency to mirrors [10,22]. Therefore, an age comparison is expected to show differences in visual scanning behavior.

In the present study, we examine a sample of 60 drivers (split equally between younger and older cohorts) as they engage in highway driving behaviors. This sample is drawn from a larger study of driver responses to the use of an in-vehicle voice system [16]. Glance behavior and task completion time metrics are analyzed during extended periods of single-task driving and during the concurrent performance of a radio tuning task using both visual-manual controls and the auditory-vocal equivalent via the SYNC system. We present statistical analyses of age cohort and task modality effects. In addition, we subdivide off-road glances for a deeper examination of glances that may be situationally-relevant or device-centric.

Methods

Participants

Study participants in two age cohorts (20-29 and 60-69) were recruited from the greater Boston area. Participants were required to possess a valid driver's license, to have had their license for more than three years, to drive on average three or more times per week, and to be in self-reported reasonably good health for their age. A research assistant verified that participants clearly understood and spoke English. Individuals were excluded if they had been involved in a police-reported accident in the past year, had a major medical illness resulting in hospitalization in the past 6 months, had a diagnosis of Parkinson's, Alzheimer's, dementia, mild cognitive impairment, or other neurological problem, were being treated for a psychological or psychiatric disorder, or had a history of cardiac disease or diabetes. Participants were informed that the expected duration of the study was four to four and a half hours, including approximately two hours of on-road driving. Compensation was $90. All participants provided informed consent, consistent with guidelines set forth by the Massachusetts Institute of Technology's Institutional Review Board.

A total of 60 participants are included in the final analysis, equally split between the younger (20-29) and older (60-69) age cohorts. Each age cohort was also split evenly by gender, resulting in 15 participants per age × gender cell. Ages did not differ significantly between genders in the younger cohort (Men $M = 24.0$ years, SD 2.7; Women $M = 24.7$ years, SD 3.0) or the older cohort (Men $M = 66.2$ years, SD 2.9; Women $M = 64.1$ years, SD 3.0)

Apparatus

Vehicle

A 2010 Lincoln MKS with factory installed voice-command systems (SYNC^{TM} for voice control of the phone and media connected by USB and the "next-generation navigation system" with Sirius Travel Link) was selected as a convenient example of a widely available production-level voice interface. The interface is engaged using a "push-to-talk" button on the right side of the steering wheel (see Figure 1). When the voice control interface is active, a display screen in the center stack supplies supporting information on system status and provides information on prompts that the driver may use in dialog with the system. A voice recognition training option is available in the system to optimize system capacity to recognize commands from an individual driver. This system training feature was utilized when a participant was introduced to the system to maximize the capacity of the system to correctly recognize commands from each participant.

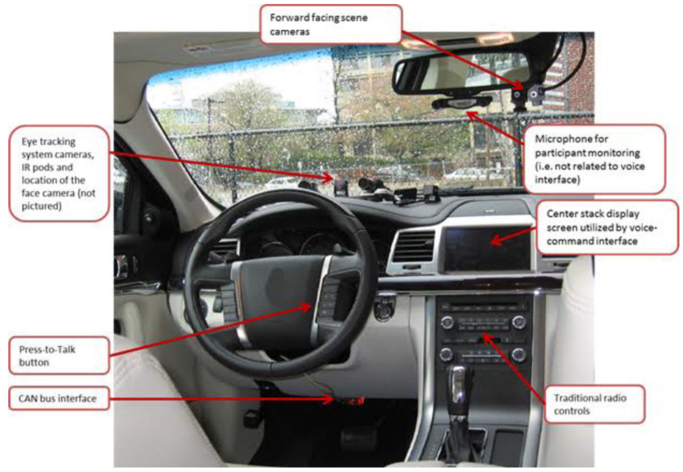

Figure 1. Photograph of the vehicle cabin and console. The press-to-talk button, which is located on the right side of the steering wheel, is used to engage the in-vehicle infotainment system's voice interface.

An experimenter was seated in the rear of the vehicle and was responsible for providing driving directions, ensuring safe vehicle operation, and verifying that participants understood and followed instructions according to a predefined script. The on-board data acquisition system supported playing recorded audio and the experimenter used a set of F-key presses at predefined points to trigger steps in the experiment. This ensured that primary instructions and tasks were presented in a consistent manner.

Cameras

The vehicle was instrumented with a customized data acquisition system that included six cameras recording simultaneously. The driver face camera, which was used to collect glance behavior, was an AVT Guppy F033C/Kowa LM6NCM with a 6mm lens, which recorded in full color at 15 frames per second.

Route

Highway driving was conducted on roadways in the greater Boston area and divided into four segments. The first segment consisted of a period of approximately 10 minutes of urban driving to reach interstate highway I-93 and continued north on I-93 for an additional 20 minutes to the I-495 intersection. This allowed a total adaptation period of approximately 30 minutes of driving prior to the assessment portion of the study. The second segment consisted of driving south on I-495 and averaged approximately 40 minutes. The third was from a rest area back north on I-495 to I-93, and the fourth was the return on I-93 south. Radio tasks were presented in a counter-balanced order during segments two and three along with other secondary tasks, with the exception that the radio- manual and radio-voice tasks were never presented in the same segment.

Baseline Driving & Secondary Tasks

Participants performed a variety of tasks as part of a larger assessment of in-vehicle interface use during highway driving [16]. This analysis specifically examined periods of single-task driving and a set of radio tuning tasks. The overall baseline driving assessment consists of 7, 2-minute periods of single task driving that were interspersed between secondary task assessments throughout the drive. During specific task performance epochs considered here, participants were asked to tune the in-vehicle radio to a specific station using both auditory-vocal and traditional visual-manual controls. Radio tuning in each modality was performed twice.

Visual-manual radio tuning consisted of four steps: turning the radio on by pressing the console's Volume button, pressing a physical Radio button, selecting the FM2 radio band using a software touchscreen button (from among AM, FM1, FM2, Sat1, Sat2, and Sat3), and finally, tuning the system to the desired station.

Auditory-vocal radio tuning also consisted of four steps: pressing a hardware "push-to-talk" button located on the steering wheel, saying "Radio," saying the desired station frequency (for example, "100.7"), and finally, confirming the selection by saying "Yes." The system prompted the driver for action at each step in this process; for example, a press of the "push-to-talk" button resulted in the auditory prompt, "Please say a command".

Data Reduction & Analysis

Glance Coding

The methodology outlined here is based on similar procedures developed by the Crash Avoidance Metrics Partnership (CAMP) Driver Workload Metrics project [1] (see also [19]). Specifically, two research assistants independently coded each radio tuning trial and single-task driving period, labeling a driver's glance targets according to one of seven categories: road, center stack, rearview mirror, instrument cluster, left window, right window, and other. Glance annotations were compared to check for discrepancies between the coders. A trial was considered discrepant if any of the following occurred: the coders started or ended their coding at different times, the coders described differing numbers of glances, the coders identified a different glance target for a glance, or the timing of a glance differed by more than 200ms. A third coder resolved any discrepancies, making a "final determination" as to which of the original two coders was correct.

Data Analysis

Following NHTSA guidelines for the assessment of visual-manual distraction, any glance off of the forward roadway was considered to be an off-road glance, and was included in the computation of glance metrics, as described below. Subsidiary analyses further subdivided off-road glances into three categories: center stack, "situationallyrelevant" (including glances to the rearview mirror, instrument cluster, right window, and left window), and "other". Glances to the center stack are taken as a proxy for glances to the radio DVI, as this region contained the traditional style visual-manual radio controls and touchscreen display which actively displayed radio status when active (Figure 1).

Several metrics of glance behavior were computed for each participant and period of interest, including the mean duration of single off-road glances, number of glances, total off-road glance time, time elapsed during task performance, and glances per minute.

NHTSA's guidelines on the assessment of visual-manual distraction state that, for an assumed sample of 24 participants, at least 21 (the 87.5th percentile) should have a mean off-road glance duration of less than 2.0s, a proportion of long-duration glances (> 2.0s) of less than 15%, and total off-road glance time of less than 12.0s. These metrics were also calculated for inclusion in analyses. Note that when these NHTSA metrics are applied for the purpose of product evaluation, they are to be computed from a sample with a different age composition than the research sample used here. Our interest was in patterns across age cohorts-and so while we report results using the NHTSA metrics, the sample upon which they are based differs from the full sample which NHTSA prescribes (NHTSA prescribes a sample of 24 participants, evenly distributed across genders and four different age groups, 18-24, 25-39, 40-54, 55+ years).

Data were analyzed in R [13] using the lme4 package for linear mixed effect modeling [2]. For the purposes of statistical reporting, model results for individual effects were converted to typical F-statistics. Secondary statistical testing was used as needed and is noted inline in the text. Single-task driving data were analyzed under a mixed-effect model that included age group (as defined by NHTSA's recommended age categories, described above) and gender as fixed effects. Subject ID was included as a random effect.

Dual-task driving with radio tuning tasks were analyzed under a model that included age group, gender, task modality, and trial number as fixed effects. Subject ID was included as a random effect. Trial number was included as a fixed effect because it was assumed that a learning or familiarity effect would be reflected in task performance.

Results

Single-Task (Baseline) Driving Periods

Mean Single Off-Road Glance Duration

Mean single off-road glance duration during single-task driving periods differed significantly by age group ($F(1, 49) = 19.41$, $p < 0.001$), as shown in Figure 2. Older drivers had consistently longer mean glance durations compared to younger drivers. Mean off-road glance duration was not significantly different across genders ($F(1, 49) = 2.14$, $p = 0.150$), or the interaction of these factors ($F(1, 49) = 0.60$, $p = 0.444$).

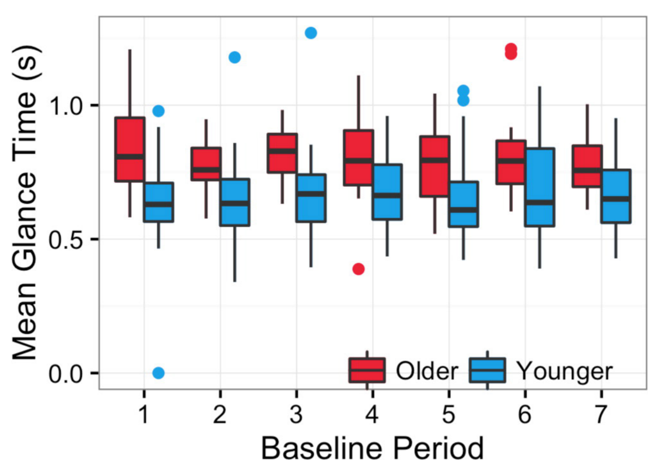

Figure 2. Mean single off-road glance duration during each of the 2-minute single-task driving periods, presented in chronological order per participant. Boxes cover the 1st and 3rd quartiles of the sample, while whiskers visualize 1.5 times beyond the interquartile range. Thick horizontal lines show group medians. Dots show outliers.

Glance Frequency (Glances per Minute)

While differences in mean glance duration by age are observed, glance frequencies are relatively similar across age groups, (Older M = 9.5 glances/min, Younger M = 8.8 glances/min; $F(1, 49) = 0.25$, $p = 0.623$). This non-significant difference in glance frequency between age groups indicates that older drivers took longer per glance than their younger counterparts, resulting in significantly elevated total off-road glance times (Older M = 15.3s, Younger M = 11.7s; $F(1, 49) = 4.84$, $p = 0.032$).

Glance Distributions

Considering subcategories of glances-situationally-relevant targets (described above), center stack (device), and other-we found that the majority of off-road glances during single task (baseline) driving were made to situationally-relevant targets (89.6% across the sample), followed by occasional glances to the center stack (7.9%, or about one per minute), and an additional 2.5% of glances were categorized as "other". Proportions of off-road glances to situationally-relevant targets were similar between age groups (88.2% for older drivers, 90.6% for younger drivers; $V = 282$, $p = 0.264$, Wilcoxon signed rank test), as were miscellaneous glances (2.1% for older drivers, 2.9% for younger drivers; $V = 315$, $p = 0.595$). There was an observed difference in the proportion of glances to the center stack (9.7% for older drivers, 6.0% for younger drivers; $V = 470$, $p = 0.025$); however, given the sparse frequency of glances to the center stack during single-task driving, this result should be interpreted with caution.

Glance Criterion Statistics

Figure 3 shows a histogram of all off-road glances made during the 14 minutes of single-task driving. The 87.5th percentile of off-road mean single glance duration, which is the criterion percentile in NHTSA's guidelines for assessing visual-manual distraction during the operation of in-vehicle electronic devices, is 0.94s for the younger cohort and 1.16s for the older cohort, less than the maximum criterion value of 2.0s.

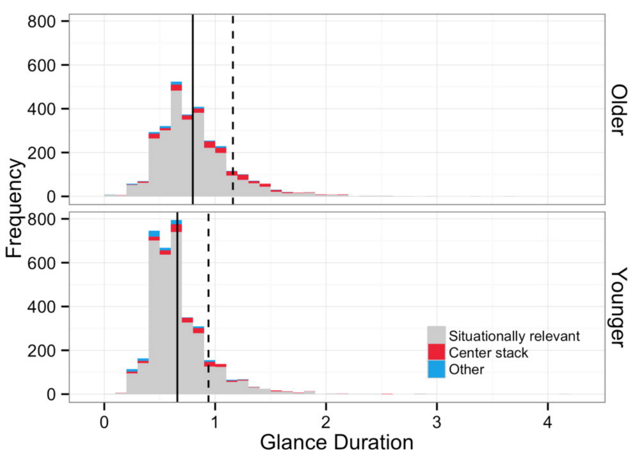

Figure 3. Histograms of off-road glance durations for each age cohort across the 14 minutes of baseline driving. Vertical solid lines represent the mean glance duration for situationally-relevant targets. Dashed lines denote the 87.5th percentile of all glances.

To examine total off-road glance time, glance times were summed per each 2-minute baseline epoch and then averaged per participant (thus each participant has one total off-road glance time metric representing a 2-minute average period). Considering any glance off the forward roadway as an "off-road" glance, the 87.5th percentiles of total off-road glance time were 17.3s for younger drivers and 22.0s for older drivers, which are greater than the criterion value of 12.0s for a "task" of undefined length. However, when situationallyrelevant glances are excluded from these calculations, criterion percentiles for total off-road glance times are reduced to 2.6s and 3.1s for younger and older drivers, respectively.

Dual-Task (Radio Tuning) Periods

Task Completion Time

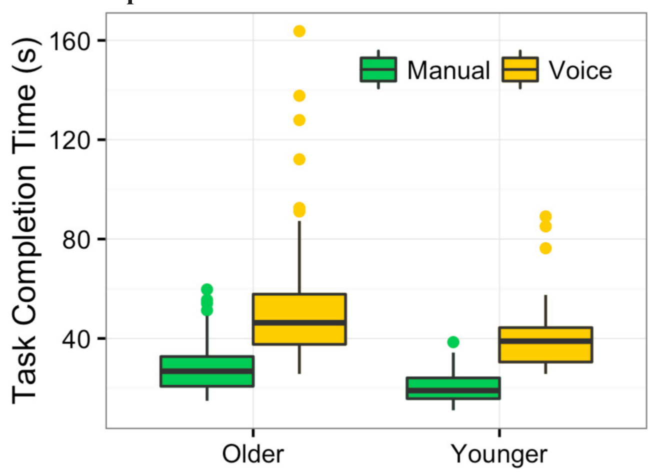

Figure 4. Task completion times during radio tuning task periods (auditory-vocal and visual-manual) for older and younger age groups. Labeling as in Figure 2.

Task completion time was significantly affected by age ($F(1, 51) = 20.64$, $p < 0.001$) and task modality ($F(1, 153) = 101.29$, $p < 0.001$). The two factors did not interact ($F(1, 153) = 1.11$, $p = 0.294$). Younger participants took 30.2s, on average, to complete the radio tasks, while older participants took 41.6s. Visual-manual radio tuning required 24.1s on average, while auditory-vocal radio tuning required 46.2s. Age and task modality did not interact significantly ($F(1, 153) = 1.11$, $p = 0.294$).

Total off-road glance time was significantly affected by age ($F(1, 49) = 28.93$, $p < 0.001$) and task modality ($F(1, 153) = 17.93$, $p < 0.001$), though the two factors did not interact ($F(1, 153) = 0.72$, $p = 0.398$). On average, younger participants spent 8.5s glancing off-road, while older participants spent 14.6s. Visual-manual radio tuning produced 13.2s of off-road glances, while auditory-vocal tuning produced 9.1s.

Task completion time was also strongly correlated with the total amount of time spent glancing off-road (Pearson's $R = 0.53$; $t(209) = 9.0$, $p < 0.001$). Therefore, to prevent task completion time from confounding measures of glance behavior, the following analyses consider only mean off-road glance duration and glances per minute, which are not directly linked to task duration.

Mean Single Off-Road Glance Duration

As shown in Figure 5, mean single off-road glance duration was significantly affected by age (Younger M = 0.79s, Older M = 0.98s; $F(1, 49) = 17.12$, $p < 0.001$) and by modality (Manual M = 0.99s, Voice M = 0.75s; $F(1, 153) = 104.52$, $p < 0.001$). These two factors did not interact significantly.

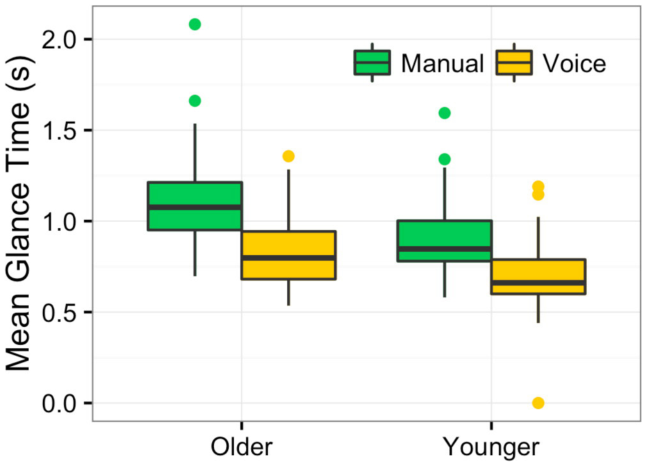

Figure 5. Mean single off-road glance duration during radio tuning task periods (auditory-vocal and visual-manual). Labeling as in Figure 2.

Glance Frequency (Glances per Minute)

Paralleling the pattern seen during single-task baseline driving periods, the rate of glances per minute was not significantly affected by age (Younger $M = 24.79$ GPM, Older $M = 23.69$ GPM; $F(1, 49) = 0.67$, $p = 0.416$). Glance rate was, however, affected by task modality; participants had an increased number of glances per minute when completing visual-manual radio tuning as compared to the auditory-vocal tuning (Manual $M = 34.08$ GPM, Voice $M = 14.46$ GPM; $F(1, 153) = 670.70$, $p < 0.001$).

Glance Distributions

Figure 6 shows glance allocation histograms for each task modality and age cohort during radio tuning (note that Figure 5 excludes glances to the forward roadway for visualization, while the percentages discussed here include them in the denominator). In contrast to the glance histograms for baseline driving, a large number of off-road glances are made to the center stack. During visual-manual task periods, younger drivers spent 50.5% of the time glancing at the center stack, while older drivers spent 56.6%. During auditory-vocal task periods, younger drivers spent 9.9% of the time glancing at the device screen, and older drivers spent 12.1%. This modality effect also extends to the allocation of situationally-relevant off-road glances: 0.9% for younger drivers performing visual-manual tuning, 2.1% for older drivers performing visual-manual tuning, 6.8% for younger drivers performing auditory-vocal tuning, and 7.9% for older drivers performing auditory-vocal tuning. Tests of mean percentage of situationally-relevant glances per subject and modality show that the greater percentage of situationally-relevant glances during voice tasks was significant ($V = 0$, $p < 0.001$, Wilcoxon signed rank test).

While the relative proportions of glance allocations among the three categories is overtly different between task modalities (as described above), there is no significant evidence to suggest that the relative proportion of glance allocations differed by age group.

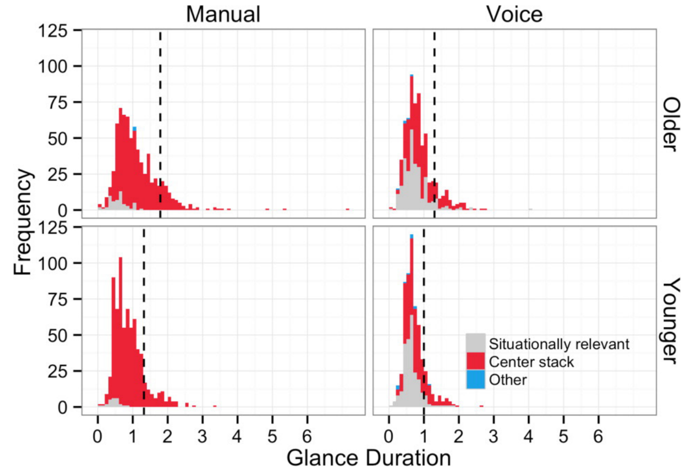

Figure 6. Histograms of off-road glance durations for each age cohort and task modality. The vertical dashed lines denote the 87.5th percentile.

Glance Criterion Statistics

Table 1 shows 87.5th percentiles for mean single off-road glance duration and total-off road glance time, as well as a total off-road glance time metric that excludes situationally-relevant targets from the calculation (in other words, including only glances to the center stack and miscellaneous "other" glances). While both radio tuning modalities in both age groups easily maintain an 87.5th percentile value of less than the 2.0s criterion, only the younger age cohort performing the auditory-vocal tuning task met the 12.0s glance time criterion if it were applied to this sample.

Table 1. 87.5th percentile of mean single glance duration, total off-road glance time, and total off-road glance time excluding situationally-relevant targets for both age groups and in all single- and dual-task periods under study. Stars (*) indicate that the calculated value would be acceptable under NHTSA guidelines for visual-manual distraction if applied to this sample and conditions.

		Mean Glance Time	Total Off-Road Glance Time	Total Off-Road Glance Time, Excluding SR Targets
Older	Single-Task Driving	0.89s*	21.95s	3.11s*
	Dual-Task Manual	1.32s*	21.47s	20.65s
	Dual-Task Vocal	1.05s*	18.11s	14.40s
Younger	Single-Task Driving	0.80s*	17.32s	2.63s*
	Dual-Task Manual	1.09s*	12.93s	12.61s
	Dual-Task Vocal	0.90s*	11.13s*	8.78s*

Summary/Conclusions

The results of the present study show that younger and older drivers exhibit some significantly different glance behaviors during highway driving. While both age cohorts make glances off the forward roadway with similar frequency, older drivers, on average, glance longer when they do so. During baseline driving, proportions of glances to situationally-relevant targets, the center stack, and a miscellaneous category were largely similar, with some evidence that the older cohort may have glanced to the center stack slightly more often than the younger cohort. Visual-manual radio tuning tasks took less time to complete than their auditory-vocal equivalents, but resulted in more time spent glancing off-road. Older drivers also took longer to complete secondary tasks than younger drivers, and spent more time glancing off-road. This resulted in a pattern that paralleled the one seen during baseline driving, with older drivers making longer off-road glances, on average, than younger drivers. Glance rates, expressed as glances per minute, were significantly higher during visual-manual task periods, but not statistically different between age cohorts. Auditory-vocal tasks also produced a significantly higher proportion of situationally-relevant off-road glances for both age cohorts.

NHTSA's guidelines on visual-manual distraction associated with the use of in-vehicle electronic devices specify that, during a task epoch, the 87.5th percentile of total off-road glance time for the sample should not exceed 12.0s. Keeping in mind that the guidelines assume the use of a driving simulator, they define an off-road glance as any glance off of the forward roadway. Applying these guidelines to an averaged 2-minute sample of single-task driving under the actual highway driving conditions in this study, the 87.5th percentile of total off-road glance time was 17.3s for younger drivers and 22.0s for older drivers, both of which are in excess of the 12.0s criterion. However, when situationally-relevant glances, including glances to vehicle mirrors and the instrument cluster, are excluded from the calculation, off-road glance times fall to 2.6s for younger drivers and 3.1s for older drivers. Conversely, total off-road glance time metrics for dual-task periods show much more modest shifts for the visual-manual task periods, regardless of the glance categorization used (note the adjustments to total off-road glance time between columns 4 & 5 in Table 1); shifts in the total glance time metric are more substantive for the voice task periods. This suggests that the NHTSA glance criteria may transfer relatively well in this respect to on-road for visual-manual task assessments. However, a "situationallyrelevant" adjustment may be worth considering, particularly for voice-command involved tasks, if it is desirable to conduct a test evaluation on the road.

Across all epochs examined in this study, older drivers had longer mean single glance durations than younger drivers, although glance frequencies were relatively similar between age cohorts. This translates into a longer total off-road glance time for the older drivers. The reasons for this consistent difference are not readily apparent in the available data, though they are in keeping with earlier work examining glance behavior [20]. It may be the case that older drivers, having greater driving experience, are simply more comfortable taking their eyes off of the forward roadway for slightly longer periods of time. On the other hand, longer glance times may reflect some combination of visual acuity loss or cognitive slowing in the older cohort. In other words, older drives may require slightly more time to extract information during a glance away from the roadway, whether to read a sign, assess a traffic pattern, or interact with a DVI. Consistent with previous findings on the comparative effects of voice interfaces over their manual equivalents, auditory-vocal tasks took longer to complete than visual-manual tasks, but resulted in less time spent glancing off-road [3, 6]. Nevertheless, as Figure 6 makes clear, drivers still spent a considerable amount of time making glances to the center console region. This contradicts the expectation that auditory-vocal interactions are "eyes-free". There is also some evidence in these data that, during single-task driving periods, older drivers glanced slightly more frequently at the in-vehicle device screen. The center console's home screen, which was active when a task associated interface was not being actively engaged, showed the driver a live map of the vehicle's current position. This live animation may have proved slightly more distracting for the older age cohort, who may be less accustomed to this type of responsive map display.

These results highlight the need for a more nuanced consideration of glance behavior during highway driving. Experimental designs and recommendations for guidelines need to consider 1) that visual behaviors may change across the lifespan 2) that glance allocation profiles may differ between simulator settings and real-world driving 3) that not all glances away from the forward roadway are necessarily "off-road" in a situationally relevant context, 4) that auditory-vocal interfaces are not inherently free of visual demand simply because they involve voice, and 5) that a dynamically-changing active display area has the potential to draw attention away from the forward roadway, even during non-task epochs.

References

1. Angell L, Auflick J, Austria P A, et al. 2006. *Driver Workload Metrics Task 2 Final Report & Appendices*. National Highway Traffic Safety Administration / US Department of Transportation, Washington, DC.
2. Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1. http://doi.org/10.18637/jss.v067.i01
3. Carter C and Graham R. 2000. Experimental Comparison of Manual and Voice Controls for the Operation of in-Vehicle Systems. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 44, 20: 3-286-3-289. http://doi.org/10.1177/154193120004402016
4. Chiang, D., Brooks, A., and Weir, D., "Comparison of Visual-Manual and Voice Interaction With Contemporary Navigation System HMIs," SAE Technical Paper 2005-01-0433, 2005, doi:10.4271/2005-01-0433.
5. Jeff Greenberg, Louis Tijerina, Reates Curry, et al. 2003. Driver Distraction: Evaluation with Event Detection Paradigm. *Transportation Research Record: Journal of the Transportation Research Board* 1843: 1-9. http://doi.org/10.3141/1843-01
6. Harbluk J.L., Burns P C, Lochner M, and Patricia L Trbovich. 2007. Using the lane-change test (LCT) to assess distraction: Tests of visual-manual and speech-based operation of navigation system interfaces.
7. Harbluk Joanne L, Noy Y Ian, Trbovich Patricia L, and Eizenman Moshe. 2007. An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. *Accident Analysis & Prevention* 39, 2: 372-379. http://doi.org/10.1016/j.aap.2006.08.013

8. Jannette Maciej and Mark Vollrath. 2009. Comparison of manual vs. speech-based interaction with in-vehicle information systems. *Accident Analysis & Prevention* 41, 5: 924-930. http://doi.org/10.1016/j.aap.2009.05.007

9. Bruce Mehler, David Kidd, Bryan Reimer, Ian Reagan, Jonathan Dobres, and Anne McCartt. 2015. Multi-modal assessment of on-road demand of voice and manual phone calling and voice navigation entry across two embedded vehicle systems. *Ergonomics*: 1-24. http://doi.org/10.1080/00140139.2015.1081412

10. Mourant R R and Rockwell T H. 1972. Strategies of visual search by novice and experienced drivers. *Human Factors* 14.

11. National Highway Traffic Safety Administration. 2013. *Visual-Manual NHTSA Driver Distraction Guidelines For In-Vehicle Electronic Devices*.

12. Owens, J., McLaughlin, S., and Sudweeks, J., "On-Road Comparison of Driving Performance Measures When Using Handheld and Voice-Control Interfaces for Mobile Phones and Portable Music Players," *SAE Int. J. Passeng. Cars - Mech. Syst.* 3(1):734-743, 2010, doi:10.4271/2010-01-1036.

13. R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. Vienna, Austria. Retrieved from http://www.R-project.org/

14. Ranney Thomas A, Baldwin G H Scott, Parmer Ed, Joshua Domeyer, Martin John, and Mazzae Elizabeth N. 2011. *Developing a Test to Measure Distraction Potential of In-Vehicle Information System Tasks in Production Vehicles*. U.S. Department of Transportation National Highway Traffic Safety Administration (NHTSA). http://doi.org/10.1037/e563342012-001

15. Ranney Thomas A, Mazzae Elizabeth N, Baldwin G H Scott, and Salaani M Kamel. 2007. *Characteristics of Voice-Based Interfaces for In-Vehicle Systems and Their Effects on Driving Performance*. Department of Transportation National Highway Traffic Safety Administration (NHTSA), Washington, DC. http://doi.org/10.1037/e729252011-001

16. Reimer Bryan, Mehler Bruce, Dobres Jonathan, and Coughlin Joseph F. 2013. *The effects of a production level "voice-command" interface on driver behavior: reported workload, physiology, visual attention, and driving performance*. Cambridge, MA.

17. Reimer Bryan, Mehler Bruce, Dobres Jonathan, et al. 2014. *Effects of an "Expert Mode" Voice Command System on Task Performance, Glance Behavior & Driver Physiology*. ACM, New York, New York, USA. http://doi.org/10.1145/2667317.2667320

18. Shutko, J., Mayer, K., Laansoo, E., and Tijerina, L., "Driver Workload Effects of Cell Phone, Music Player, and Text Messaging Tasks with the Ford SYNC Voice Interface versus Handheld Visual-Manual Interfaces," SAE Technical Paper 2009-01-0786, 2009, doi:10.4271/2009-01-0786.

19. Smith David L, Chang James, Glassco Richard, Foley James, and Cohen Daniel. 2005. Methodology for capturing driver eye glance behavior during in-vehicle secondary tasks. *Transportation Research Record: Journal of the Transportation Research Board* 1937, 1: 61-65.

20. Tijerina L and Parmer E. 1998. Driver workload assessment of route guidance system destination entry while driving: A test track study. Proceedings of the 5th ITS World Congress, Seoul, South Korea.

21. Victor Trent W, Harbluk Joanne L, and Engström Johan A. 2005. Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behaviour* 8, 2: 167-190. http://doi.org/10.1016/j.trf.2005.04.014

22. Wang Ying, Reimer Bryan, Dobres Jonathan, and Mehler Bruce. 2014. The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation Research Part F: Psychology and Behaviour* 26, PA: 227-237. http://doi.org/10.1016/j.trf.2014.08.003

23. Yager Christine. 2013. *An Evaluation of The effectivness of voice-to-text programs at reducing incidences of distracted driving*. Texas A&M Transportation Institute, College Station, TX.

24. Young Kristie L, Salmon Paul M, and Cornelissen Miranda. 2013. Missing links? The effects of distraction on driver situation awareness. *Safety Science* 56: 36-43. http://doi.org/10.1016/j.ssci.2012.11.004

Contact Information

The corresponding author of this work can be reached via jdobres@mit.edu

Acknowledgments

Support for this work was provided in part by Toyota's collaborative Safety Research Center (CSRC) and the US DOT's Region I New England University Transportation Center at MIT. Data was drawn from studies also supported by The Santos Family Foundation.

We would also like to acknowledge the contributions of research assistants Hale McAnulty, Daniel Munger, and Alea Mehler, who oversaw the intake and on-road assessment of research participants throughout the data collection phase of this study.

The Engineering Meetings Board has approved this paper for publication. It has successfully completed SAE's peer review process under the supervision of the session organizer. The process requires a minimum of three (3) reviews by industry experts.

ISSN 0148-7191

http://papers.sae.org/2016-01-1445