Scrolling and Driving: How an MP3 Player and Its Aftermarket Controller Affect Driving Performance and Visual Behavior

John D. Lee, Shannon C. Roberts, Joshua D. Hoffman and Linda S. Angell

DOI: 10.1177/0018720811429562

The online version of this article can be found at:
http://hfs.sagepub.com/content/54/2/250
Objective: The aim of this study was to assess how scrolling through playlists on an MP3 player or its aftermarket controller affects driving performance and to examine how drivers adapt device use to driving demands.

Background: Drivers use increasingly complex infotainment devices that can undermine driving performance. The goal activation hypothesis suggests that drivers might fail to compensate for these demands, particularly with long tasks and large search set sizes.

Method: A total of 50 participants searched for songs in playlists of varying lengths using either an MP3 player or an aftermarket controller while negotiating road segments with traffic and construction in a medium-fidelity driving simulator.

Results: Searching through long playlists (580 songs) resulted in poor driving performance and required more long glances (longer than 2 s) to the device compared with other playlist lengths. The aftermarket controller also led to more long glances compared with the MP3 player. Drivers did not adequately adapt their behavior to roadway demand, as evident in their degraded driving performance. No significant performance differences were found between short playlists, the radio-tuning task, and the no-task condition.

Conclusion: Selecting songs from long playlists undermined driving performance, and drivers did not sufficiently adapt their use of the device to the roadway demands, consistent with the goal activation hypothesis. The aftermarket controller degraded rather than enhanced performance.

Application: Infotainment systems should support drivers in managing distraction. Aftermarket controllers can have the unintended effect of making devices carried into the car less compatible with driving. These results can motivate development of new interfaces as alternatives to scrolling lists.

Keywords: driver adaptation, driver behavior, driver distraction, interface evaluation, usability, MP3 player

INTRODUCTION

A growing number of technologies promise to make driving more pleasant and more efficient. Many of these technologies, such as cell phones and MP3 players, can also distract drivers. Engaging in activities that draw attention away from driving can increase crash likelihood (Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006), and in 2008, driver distraction was reported in 16% of all fatal crashes, contributing to 5,870 deaths (Ascone, Lindsey, & Varghese, 2009). Sources of driver distraction are diverse and include eating, talking, and interacting with infotainment devices (Regan, Lee, & Young, 2008; Stutts, Reinfurt, Staplin, & Rodgman, 2001).

The diversity of infotainment devices presents a particular challenge. Broadly considered, infotainment devices fall into three categories: those integrated into vehicles by manufacturers (e.g., radios and in-vehicle navigation systems), those carried into vehicles but not designed for use in vehicles (e.g., cell phones and MP3 players), and those that adapt carried-in devices to vehicles (e.g., docking stations and supplementary controllers). Adapters are designed to make devices less distracting and more compatible with the demands of driving, but there is no testing required to verify these benefits.

Of particular interest is the MP3 player. The desire to carry MP3 players into vehicles is so widespread that 90% of new vehicles sold in the United States have an option for iPod connectivity (Apple, 2011). In a survey of 5,288 drivers ages 18 to 24, 20% reported selecting songs on their MP3 players while driving (GMAC Insurance, 2006). In addition, younger drivers tend to engage in more risky driving behavior and are not as efficient in processing critical driving information (Wikman, Nieminen, &
Summala, 1998). Given these findings, it is particularly important to investigate how MP3 player use affects young drivers.

Because MP3 players are operated in vehicles, it is important to understand the attentional demands they place on drivers and to understand how MP3 player adapters might alter these demands. Devices with highly demanding visual-manual interfaces, such as MP3 players, and those that impose high cognitive demands can undermine driving performance (Hoffman, Lee, McGeehe, Macias, & Gellatly, 2005; Horrey & Wickens, 2006; Strayer, Drews, & Johnston, 2003; Tsimhoni, Smith, & Green, 2004). The cognitive demand associated with complex MP3 tasks, such as searching through playlists, might compound the visual demand of reading and selecting from a playlist, leading to long glances away from the road and degraded performance.

In four previous simulator studies, MP3 player interaction was shown to undermine driving performance. In one study, researchers compared an MP3 player with two interfaces that operate the MP3 player through voice commands (Garay-Vega et al., 2010). The visual-manual interaction with an MP3 player enabled drivers to find music quicker, but caused drivers to spend more time with their eyes off the road, compared with interaction through voice commands. In another study, researchers found that tasks that required significant manipulation of the MP3 player (i.e., skipping song tracks and changing the volume) resulted in significant decrements in lane-keeping performance and increases in lateral and longitudinal speed variability, as compared with an in-person or cell phone conversation (Crisler et al., 2008). Selecting artists, songs, videos, and podcasts on an MP3 player while driving increased lateral deviation from the lane center and was associated with slower driving relative to baseline driving (Salvucci, Markley, Zuber, & Brumby, 2007). List scrolling had a greater effect on performance compared with selecting items from a menu.

A similar study addressed how MP3 player interaction affected response to pedestrians, vehicles entering the roadway, and braking vehicles (Chisholm, Caird, Lockhart, Fern, & Teteris, 2007). Scrolling through long playlists (900 songs) led to longer event response times and more glances away from the road compared with single button presses. Drivers also looked away from the road more when speed limits and traffic decreased, suggesting that they strategically balanced roadway demands with those of the MP3 player.

By strategically modulating their use of MP3 players in response to changing roadway demands, drivers might reduce the conflict between the roadway and MP3 player (Lee, Regan, & Young, 2008); however, long tasks might undermine this strategic behavior. According to the goal activation model, goals with higher activation are more likely to be attended (Altmann & Trafton, 2002). Goals with high activation include those that are of high importance or those that are close to completion. Activation of an unattended goal decays over time, and reengaging this unattended goal often depends on activation-enhanced priming from the task environment. For example, activation for the goal of searching through a playlist to find a song might grow after it is initiated, leading to neglect of the driving task and to resumption delays for driving that increase with MP3 task duration. Goal activation for the search task could undermine drivers’ ability to strategically modulate attention and cause long search tasks to be distracting. Therefore, the goal activation model predicts that playlist length, and more generally task length, would be a critical design variable that influences drivers’ strategic modulation of attention between driving and a secondary task. The goal activation model provides a theoretical framework to predict how task length might influence driving performance.

This study had three goals: (a) to assess how playlist length governs the interference of scrolling with driving, (b) to investigate how an aftermarket controller that adapts MP3 players to vehicles affects driving performance associated with scrolling through playlists, and (c) to determine whether drivers modulate their attention to the MP3 player to accommodate changes in roadway demand.

**METHOD**

**Participants**

For this study, 50 young drivers (29 males, 21 females) ages 18 to 25 (\(M = 21, SD = 2\))
participated in the experiment. Each participant had a minimum of 2 years of driving experience ($M = 5, \ SD = 2$) and had normal to corrected-to-normal vision. All participants had owned an MP3 player for at least 4 months ($M = 18$) and used it while driving at least once a week ($M = six \ times \ per \ week$).

**Apparatus**

This study used a fixed-based, medium-fidelity simulator based on a full 1992 Mercury Sable vehicle cab that had standard controls, including brake, throttle, and a force-feedback steering wheel, and included a horizontal field of view of approximately $50^\circ$. Driver eye movements were collected with the use of Seeing Machines’ faceLAB eye tracker (Version 4.2), which uses two small cameras mounted on the dash of the simulator to track head movements and eye movements (Figure 1).

**Driving Environment**

Participants drove on a straight, five-lane suburban roadway that included two lanes in each direction separated by a center turning lane. Participants were instructed to drive at 45 mph (72.4 km/h), stay in the center of the right-hand lane, and make no turns. The roadway included alternating sections of construction, traffic, and baseline roadways (Figure 2). Construction sections consisted of left-lane closures indicated by construction barrels that encroached 0.5 m into the right-hand lane and reduced the lane width from 3.66 m to 3.16 m. Traffic sections required drivers to follow a lead vehicle closely and drive next to vehicles traveling approximately the same speed in the adjacent lane. The traffic section required vigilant longitudinal and lateral control to avoid a collision. A following vehicle responded to large time headways (greater than 1.5 s) between the driver and lead vehicle by honking, prompting drivers to maintain a consistent following distance. Baseline driving refers to segments that had neither the congestion nor traffic demands. A traffic light separated each roadway section.

Drivers were required to press a button on the steering wheel when they detected bicyclists on the roadway. Bicyclists were randomly distributed so that the participant passed two to three bicyclists per minute for a total of 44 bicyclists during the drive. Because bicyclists were not uniformly distributed across tasks for all drivers (because of the time differences required to complete tasks), the detection of bicyclists was analyzed across entire roadway sections (baseline, construction, and traffic). Drivers also responded to a braking lead vehicle while performing a task with traffic, a task with no traffic, and no task with no traffic. There were no braking lead vehicles in the construction section.

**Search Tasks**

For each search task, a synthetic voice instructed participants to engage in the task at specified points during the drive (Figure 2). The MP3 player search tasks required drivers to scroll through a short (20 songs), medium (75 songs), or long (580 songs) playlists sorted alphabetically by song title to find a target song. Drivers had to scroll through 7 to 12 songs in the short playlist, 21 to 30 songs in the medium playlist, and 100 to 130 songs in the long playlist to find the target song. These lengths were chosen to ensure that the short playlist tasks required little to no scrolling, the medium playlist task required one full scroll, and the long playlist task required multiple scrolls. Drivers responded verbally with the song number in the playlist (which was displayed) to indicate task
completion. For each search task, drivers were required to find a different target song; that is, drivers did not search for the same song twice.

As a point of comparison, all drivers performed a radio tuning task. Following Angell et al. (2006), drivers tuned the radio to a particular frequency. Frequencies alternated between the AM and FM bands, and the targets were located 23 to 35 increments (0.2 Hz for FM, 10 Hz for AM) from their current location. Drivers responded by saying “Done” when the station was found. Search tasks were presented before and within traffic and construction sections as well as during baseline periods (Figure 2).

**Devices**

A 30-GB, fifth-generation iPod with color screen was mounted in a cradle on the center instrument stack. The MP3 player featured momentum scrolling; that is, the scrolling speed is dictated by how fast the user gestures. An aftermarket controller designed for the MP3 player (Harman/Kardon Drive + Play) displayed menu information on a small screen and had a separate controller that duplicated the functionality of the MP3 player (Figure 1). The controller featured proportional scrolling; that is, when the user commanded maximum scrolling, the playlist was advanced at a rate proportional to the playlist length. The aftermarket display was mounted on the dash in the same location as the MP3 player, and the controller was located at the end of the participants’ reach when their right arm was resting at their side.

**Experimental Design**

The experiment used a 2 (interface) × 5 (in-vehicle task) × 3 (roadway section) × 3 (demand timing) mixed design. Interface was the only between-subjects variable and consisted of two levels: MP3 player or the aftermarket controller. The in-vehicle task had five levels: no task, MP3 song selection from short, medium, or long playlist; and the radio tuning task. Roadway demand consisted of baseline, traffic, and construction sections. Both in-vehicle task and roadway demand were counterbalanced across participants. Demand timing defines the search task request relative to the demanding sections of the roadway. Demand timing had three levels: baseline (i.e., no traffic or construction), prior to demand (i.e., before entering traffic or construction), and during demanding roadway sections (i.e., during traffic or construction).

**Procedure**

After arrival, participants gave consent and completed a MP3 player-use questionnaire. Next, they drove a 5-min practice route to become familiar with the simulator dynamics and roadway demands as well as the radio tuning and medium playlist search task. Participants using the aftermarket controller were given extra time before the practice route to become familiar with the device and did not continue with the study until they felt comfortable using the device while driving. Participants drove five times, each corresponding to an in-vehicle task and a no-task drive, and were presented with the same search task 16 times during each drive (Figure 2).

Before each drive, participants received instructions specific to the drive but were generally instructed to “drive as you normally would, obeying all traffic laws and signals” and...
“your task is to use the iPod controls to select the appropriate song.”

Following each drive, participants rated their workload using the NASA Task Load Index (Hart & Staveland, 1988), results of which are not included in this article. After the fifth drive, participants completed a postexperiment survey.

**Dependent Measures**

Measures computed to assess driving performance included standard deviation of lane position, speed variability (root mean square speed), and bicycle detection (d’, sensitivity). Task initiation time and eye movement metrics (number of glances and glance duration) were also measured. All measures were calculated for the roadway sections associated with each of the 16 periods in which the in-vehicle tasks were performed (Figure 2). Driving performance measures were also calculated in the no-task drive whereby each of the in-vehicle tasks was performed as matched by roadway location.

**RESULTS**

A mixed linear model with participant as a repeated measure was used, and post hoc comparisons were performed using SAS 9.1 and R 2.11.1. An Interface (two levels) × In-Vehicle Task (five levels) × Demand Timing (three levels) ANOVA model was analyzed for each of the dependent variables. The traffic and construction sections were analyzed separately, given that the two sections imposed substantially different demands. All post hoc comparisons via t tests are shown in Table 1.

**Standard Deviation of Lane Position**

During construction sections, drivers using the MP3 player demonstrated similar variability in lane position as those using the aftermarket controls, \(F(1, 48) = 0.02, p = .88\) (Figure 3). A slight increase in lane position variability was observed when long playlist tasks were performed, \(F(4, 192) = 4.45, p < .01\); otherwise, standard deviation of lane position was similar for no task, radio tuning, short playlist search, and medium playlist search. During medium and long search tasks, lane position variability was lower when the tasks were introduced before the construction section, in part because the latter portion was completed in the construction section: Task × Demand Timing, \(F(8, 381) = 2.92, p < .01\).

In traffic periods, lane position variability was not affected by the interface, \(F(1, 48) = 0.02, p = .89\). The greatest lane position variability occurred during the long playlist task: Task, \(F(4, 191) = 17.71, p < .01\). Traffic reduced the lane position variability during task performance, \(F(2, 96) = 37.71, p < .01\), compared with baseline periods.

**Speed Variability**

During construction, no effect of interface was observed, but task type and demand timing both affected speed variability, \(F(1, 48) = 1.68, p = .20\); \(F(4, 192) = 23.31, p < .01\); \(F(2, 96) = 35.15, p < .01\), respectively (MP3 player, \(M = 0.40\); aftermarket, \(M = 0.33\)). Greater speed variability was observed during long playlist tasks compared with other tasks and the no-task condition (long, \(M = 0.53\); medium, \(M = 0.36\); short, \(M = 0.31\); radio, \(M = 0.32\); no task, \(M = 0.32\)), none of which differed. Demand timing also affected speed variability, as speed was less varied during construction, compared with prior to construction or during the baseline (during, \(M = 0.27\); prior, \(M = 0.46\); baseline, \(M = 0.37\)).

Speed variability in traffic was sensitive to task type and demand timing, \(F(4, 191) = 51.85, p < .01\); \(F(2, 96) = 61.06, p < .01\), respectively (long, \(M = 0.75\); medium, \(M = 0.46\); short, \(M = 0.41\); radio, \(M = 0.41\); no task, \(M = 0.44\); baseline, \(M = 0.37\); prior, \(M = 0.59\); during, \(M = 0.52\)). A marginal effect of interface did exist, \(F(1, 48) = 3.4, p = .07\) (MP3 player, \(M = 0.45\); aftermarket, \(M = 0.54\)); however, the difference is better explained by the interaction of task and interface, \(F(4, 191) = 3.95, p < .01\). Performing long tasks with the aftermarket controller resulted in the greatest variation in speed. Task type also interacted with demand timing such that there was greater speed variability approaching and during traffic sections when drivers performed long tasks, \(F(8, 380) = 2.2, p = .03\).
Bicycle detection Sensitivity (d’) to the presence of a bicyclist depended on task type, $F(4, 191) = 15.34$, $p < .01$, but not interface, $F(1, 48) = 0.43$, $p = .52$. Drivers were less sensitive to bicyclists when performing long tasks ($M = 2.61$) compared with any other tasks (medium, $M = 2.80$; short, $M = 2.93$; radio, $M = 3.06$) or no task ($M = 3.12$). Drivers showed no difference in bicyclist detection when performing the short playlist and radio tuning tasks but were less sensitive during the short playlist task compared with no task, whereas the radio task was not different from the no-task condition.

Task Initiation Time

While in the construction section, drivers did not hesitate to start a task following its presentation, averaging 1.05 s across all conditions. Drivers initiated the task quickly regardless of interface, $F(1, 48) = 0.11$, $p = .75$, or demand timing, $F(2, 96) = 1.81$, $p = .17$. Similarly, the initiation times did not differ between tasks, although a marginal trend suggests drivers may have waited longer to begin long playlist tasks, $F(3, 144) = 2.4$, $p = .07$.

Drivers encountering traffic were quicker to begin the task compared with during the construction section, averaging 0.96 s across all conditions.

### Table 1: Statistical Results of t Tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>Effect</th>
<th>df</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard deviation of lane position during traffic</td>
<td>Prior to demand vs. baseline</td>
<td>96</td>
<td>6.81</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>During demand vs. baseline</td>
<td>96</td>
<td>7.35</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td>Speed variability during construction</td>
<td>Long vs. medium</td>
<td>192</td>
<td>6.37</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. short</td>
<td>192</td>
<td>8.18</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. radio</td>
<td>192</td>
<td>7.83</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. no task</td>
<td>192</td>
<td>7.56</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Prior vs. during demand</td>
<td>96</td>
<td>8.38</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Prior to demand vs. baseline</td>
<td>96</td>
<td>4.67</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td>Bicycle detection</td>
<td>Long vs. medium</td>
<td>191</td>
<td>-2.57</td>
<td>.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. short</td>
<td>191</td>
<td>-4.42</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. radio</td>
<td>191</td>
<td>-5.89</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Short vs. radio</td>
<td>191</td>
<td>1.46</td>
<td>.15</td>
</tr>
<tr>
<td></td>
<td>Short vs. no task</td>
<td>191</td>
<td>2.59</td>
<td>.01*</td>
</tr>
<tr>
<td></td>
<td>Radio vs. no task</td>
<td>191</td>
<td>1.15</td>
<td>.25</td>
</tr>
<tr>
<td>Task initiation time during traffic</td>
<td>Radio vs. short</td>
<td>144</td>
<td>2.39</td>
<td>.02*</td>
</tr>
<tr>
<td></td>
<td>Medium vs. short</td>
<td>144</td>
<td>3.47</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. short</td>
<td>144</td>
<td>3.61</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td>Number of glances to device during construction</td>
<td>Long vs. medium</td>
<td>140</td>
<td>22.83</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. short</td>
<td>140</td>
<td>28.90</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. radio</td>
<td>140</td>
<td>27.96</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Short vs. radio</td>
<td>140</td>
<td>1.23</td>
<td>.22</td>
</tr>
<tr>
<td>Number of glances to device during traffic</td>
<td>Long vs. medium</td>
<td>141</td>
<td>18.95</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. short</td>
<td>141</td>
<td>25.99</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Long vs. radio</td>
<td>141</td>
<td>25.10</td>
<td>&lt;.01*</td>
</tr>
<tr>
<td></td>
<td>Short vs. radio</td>
<td>141</td>
<td>1.23</td>
<td>.22</td>
</tr>
<tr>
<td></td>
<td>Prior vs. during demand</td>
<td>96</td>
<td>2.03</td>
<td>.05*</td>
</tr>
<tr>
<td></td>
<td>Prior to demand vs. baseline</td>
<td>96</td>
<td>4.34</td>
<td>&lt;.01*</td>
</tr>
</tbody>
</table>

*p ≤ .05.
conditions. There was a main effect of task, $F(3, 144) = 5.57, p < .01$, with drivers waiting longer to begin the radio tuning and medium and long playlist tasks than the short playlist task. Similar to construction, there was no effect of interface, $F(1, 48) = 0.01, p = .92$, or demand timing, $F(2, 96) = 0.7, p = .50$.

**Number of Glances to the Device**

The number of glances required to perform the task in construction sections was sensitive to task type, $F(3, 140) = 368.92, p < .01$; interface, $F(1, 48) = 18.16, p < .01$; and demand timing, $F(2, 96) = 41.32, p < .01$. Drivers using the aftermarket controller required an average of two additional glances ($M = 8.21$) to complete the tasks compared with those using the MP3 player ($M = 6.31$). Long playlist tasks required twice as many glances ($M = 12.09$), compared with medium playlist ($M = 6.64$), short playlist ($M = 5.00$), and radio tuning ($M = 5.31$). No statistically significant difference existed between the short playlist and radio tuning task.

In traffic sections, task type, $F(3, 141) = 290.18, p < .01$ interface, $F(1, 48) = 17.89, p < .01$; the Task × Interface interaction, $F(3, 141) = 9.19, p < .01$; and demand timing, $F(2, 96) = 9.56, p < .01$, affected the number of glances. The long playlist task required an average of 10.5 glances, substantially more than did other tasks (medium, $M = 6.37$; short, $M = 4.76$; radio, $M = 5.03$). The short playlist and radio

---

**Figure 3.** Standard deviation of lane position across tasks and demand timing. B is the baseline period (between traffic and construction sections), P is predemand, and D is during the demanding section of the roadway.

---

- **Construction**
  - Standard Deviation of Lane Position (m)
  - **Standard Deviation of Lane Position (m)**
  - **Standard Deviation of Lane Position (m)**
  - **Standard Deviation of Lane Position (m)**
  - **Standard Deviation of Lane Position (m)**

- **Traffic**
  - Standard Deviation of Lane Position (m)
  - **Standard Deviation of Lane Position (m)**
  - **Standard Deviation of Lane Position (m)**
  - **Standard Deviation of Lane Position (m)**
  - **Standard Deviation of Lane Position (m)**
tuning task were similar. Using the aftermarket controller resulted in a greater number of glances compared with using the MP3 player (MP3 player, $M = 5.89$; aftermarket: $M = 7.42$), and this difference increased with playlist length. Tasks performed prior to the traffic section required more glances ($M = 7.05$) compared with during traffic ($M = 6.64$) or during the baseline ($M = 6.28$).

**Glance Duration to the Device**

During construction sections, glances to the long playlist were more than 160 ms longer than for other tasks, $F(3, 140) = 21.48, p < .01$. With the aftermarket controller, glance durations averaged 127 ms longer compared with the MP3 player, $F(1, 48) = 4.13, p = .05$. Specifically, longer glance durations were seen for drivers performing the long playlist task using the aftermarket controller compared with any of the other tasks or using the MP3 player: Interface $\times$ Task, $F(3, 140) = 6.09, p < .01$. Glance duration also depended on demand timing, $F(2, 96) = 40.25, p < .01$, and the interaction between task type and demand timing, $F(6, 275) = 2.27, p = .04$. Glance duration decreased when construction demands coincide with the long playlist tasks, particularly when drivers performed tasks prior to construction sections. Drivers did not compensate for longer glances to the device by looking at the roadway longer while interleaving tasks, $F(3, 144) = 0.55, p = .65$.

During traffic sections, long playlist tasks led to longer glances, $F(3, 141) = 25.86, p < .01$, as did using the aftermarket controller, $F(1, 48) = 5.67, p = .02$. Task type and interface interacted such that glance duration increased by almost 225 ms when the aftermarket controller was used to perform the medium and long playlist tasks compared with when the MP3 player was used, $F(3, 141) = 4.95, p < .01$. Glance durations decreased slightly as traffic demands increased, $F(2, 96) = 4.16, p = .02$. As in the construction section, drivers did not compensate for longer glances to the device by looking at the roadway longer between device interactions, $F(3, 144) = 1.83, p = .14$.

Figure 4 shows a histogram of glance duration by interface and task, and Figure 5 shows the density function. In Figure 5, the dark solid line indicates the 95th percentile glance duration for the first 15 s of the task, and the gray solid line indicates the 95th percentile glance duration for the last 15 s of the task. Glances occurring later in the task tend to be longer, but for the MP3 player during the traffic section, the 95th percentile glance duration is shorter for the last 15 s of the task. The aftermarket controller required more glances longer than 2 s and more long glances toward the end of the task compared with the MP3 player. Following the 15-s rule established by SAE J2634 (Society of Automotive Engineers, 2004), glance occurrence was divided into those that occurred within the first 15 s of the task and those that occurred after the first 15 s. The mean glance duration during the first 15 s of the task (construction, $M = 1.09$; traffic, $M = 1.14$) was shorter than the glance duration after the first 15 s (construction, $M = 1.43$; traffic, $M = 1.47$) in both the construction and traffic sections: construction, $F(1, 49) = 95.5, p < .01$; traffic, $F(1, 49) = 125.74, p < .01$. In addition, the 95th percentile glance duration for glances in the first 15 s of the task (construction, 2.47; traffic, 2.63) was shorter than for the second 15 s of the task (construction, 3.13; traffic, 3.35). Note that as the radio and short playlist task were often completed quickly (i.e., before 15 s), an analysis by task type as separated by the 15-s rule is not feasible.

Figure 6 shows that glance duration depends not only on task duration but also on the glance history—a long glance is more likely to follow another long glance. In Figure 6, both the current glance and the previous glance are transformed to logarithmic scale. The solid line acts as a reference at 2 s, the dashed line shows the 95th percentile glance duration, and the dotted line shows the mean glance duration. The mean and 95th percentile glance duration are substantially greater for the long task, and the regression line shows how likely long glances are to follow other long glances. This effect is strongest for the longer tasks and weaker for the shorter tasks: For long tasks, the regression line predicts glance durations well above the mean when glances are preceded by a long glance. For short tasks, preceding long glances predict...
glance durations that are only slightly above the mean duration.

These effects were tested with a multilevel linear model that assessed the influence of previous glance duration, task, and device on current glance duration with participant as a random factor. The statistical significance of these variables was tested by assessing whether models that included these variables fit the data better than models without, starting with a simple model that included only an intercept with drivers as a random variable. A model that included the previous glance duration predicted current glance duration much better than the model including only the intercept, with an Akaike Information Criterion (AIC) of 27,955, $\chi^2(1) = 5953, p < .0001$.

Including the interaction of task type and previous glance duration accounted for the data better than one that included only task duration, with an AIC of 27,897, $\chi^2(3) = 63, p < .01$. Similarly, as shown in Figure 6, adding the main effect of task and device to the model also improved the model fit, with the addition of task producing an AIC of 27,594, $\chi^2(3) = 309, p < .01$, and the addition of interface producing an AIC of 27,588, $\chi^2(1) = 7.9, p < .01$.

**DISCUSSION**

This study extends previous research examining MP3 player use while driving. It provides further evidence that scrolling through play-lists, particularly, long play-lists, undermines driving performance and that aftermarket devices do not necessarily mitigate this effect.

Most performance measures did not significantly differ between the radio and short play-list tasks, suggesting that the demands associated with searching through a short play-list and with radio tuning are similar. However, performance decrements were observed when drivers searched through the long play-lists: Drivers made more glances and longer glances to the device and also showed greater variability in
speed and lane position. This is consistent with previous research that found that simple MP3 player interactions had little effect (Crisler et al., 2008; Salvucci et al., 2007). When searching through long playlists, drivers also showed a decreased ability to detect bicyclists. Glance duration was also substantially longer after the first 15 s of a task had elapsed and when a glance was preceded by a long glance. These results suggest that drivers neglected the driving task as they increased their attention to the MP3 task.

This study also extends previous research by comparing an MP3 player that is carried into the vehicle with an aftermarket device designed to adapt the MP3 player to the vehicle. For example, the aftermarket controller had a relatively small effect on lane position and speed variability, but a substantial negative effect in glance frequency and duration. This result is surprising, considering that aftermarket controllers are advertised as devices to help mitigate the demands of the device and promote safe driving. In particular, it is suspected that the proportional scrolling of the aftermarket device led to poor performance, as compared with the momentum scrolling of the MP3 player. These results demonstrate that simple adaptations that have intuitive appeal may fail to provide expected benefits if detailed analyses of control dynamics and feedback are not considered.

Task initiation times indicated that drivers did not hesitate to start a short or medium-length task, regardless of roadway demand or interface, although they did hesitate to begin the long playlist task. In addition, when drivers were faced with a prolonged secondary task, their ability to adapt to the roadway demands diminished: They showed an increased variability in speed and lane position and a reduced ability to detect bicyclists. Taken together, although drivers hesitated slightly when initiating demanding tasks and took shorter
glances away from the road during demanding roadway sections, drivers were not completely successful in adapting their use of the MP3 player according to roadway conditions, as demonstrated by poorer driving performance and longer glances with longer tasks.

Such results are consistent with the goal activation model of task performance, in which activation associated with the nondriving task increases over time and activation of the driving task decays over time (Altmann & Trafton, 2002). Consistent with the goal activation model, the activation of the MP3 goal increased with playlist length, and long glances tended to follow other long glances. In addition, given that the driver was looking at the roadway, bicyclists were not a sufficient external cue to prime the driving task. The results also show evidence of a priming effect in that the presence of traffic diminishes the effect of long playlist tasks; that is, drivers took shorter glances toward the device as they approached the demanding roadway sections. This finding suggests that cues associated with traffic encroaching on the driver may enhance activation of the driving task, bringing it back into the focus of attention and diminishing attention to the search task. Overall, the results suggest that the dynamics of attention and the associated influence of goal activation could have important implications for understanding and reducing driver distraction.

Although the goal activation model is consistent with the findings of this study, particularly, the effects of task duration and the

\[ \text{Figure 6. Current glance duration as a function of previous glance duration for task type and interface. The solid horizontal line represents 2 s, the dashed line shows the 95th percentile glance duration for each condition, and the dotted line shows the mean duration for each condition. The regression line is bounded by the 95\% confidence limit.} \]
previous glance duration, this interpretation rests on several assumptions that merit further exploration. For one, the goal activation model typically applies to discrete tasks that have well-defined goals. One could argue that driving is a continuous tracking task without such goals—there is no driving-specific goal with a corresponding activation that influences performance. However, driving is actually a complex, multifaceted activity that blends elements of continuous tasks with those that are more similar to those typically considered by the goal activation model, such as hazard detection and navigation. Another assumption is that the playlist tasks are similar from beginning to end, with the negative effects seen at the end of the task attributed to goal activation of the scrolling task. In this study, drivers start scrolling to reach the general vicinity of the song, whereas at the end, drivers scrolled for the specific song. These components might pose different demands on the driver beyond the effect of goal activation.

More generally, the goal activation model typically describes the consequence of switching between tasks, whereby each task is performed for a relatively long time compared with the rapid interleaving of small chunks of tasks seen in driving. Although inconclusive, the current experiment shows that goal activation might also play a role in how well drivers interleave driving and nondriving tasks. In future research, investigators might consider a more detailed analysis of how specific interactions during the task contribute to the neglect of the driving task; however, this study shows that the goal activation consequences of long tasks merit particular consideration. Further research should include the degree to which goal activation influences various components of the driving task with measures, such as resumption lag, that can assess this influence more precisely (Monk, Boehm-Davis, & Trafton, 2004).

Generalizing this study should be done with caution. As with any simulator study, certain elements are not representative of actual driving and device use. Further investigation should include more representative samples of such factors as device mounting location and experience using the device. Mounting the MP3 player on the instrument panel might not be representative of how drivers actually use MP3 players. The mounting location might have imposed greater demands on drivers accustomed to placing the device elsewhere, such as on the steering wheel. Although participants were given time to become familiar with the aftermarket controller beforehand, this familiarity does not compare with their experience with the MP3 player (at least 6 months). Therefore, the effect of interface could be partially attributable to drivers’ greater familiarity with the MP3 player. However, drivers are not likely to practice extensively before they use an aftermarket interface while driving, and so the conditions tested in this study are representative of at least a portion of drivers’ experience with such systems.

**CONCLUSIONS AND IMPLICATIONS**

Consistent with the goal activation model, long tasks led drivers to attend less to driving and to drive poorly. Therefore, minimizing long and complicated tasks merits attention in design. Scrolling through long lists, such as a playlist, seems incompatible with the demands of driving, suggesting that new methods that are faster and less demanding are needed. Aftermarket controllers can magnify rather than diminish these demands. Because manufacturers of aftermarket controllers might inadvertently imply that such interactions are consistent with the demands of driving, other approaches to reduce distraction merit consideration. For example, one such approach could help drivers recognize the distraction posed by such devices and enable drivers to manage distraction by providing cues and alerts when they look away from the road for too long (Donmez, Boyle, & Lee, 2008; Lee, 2009).

**ACKNOWLEDGMENTS**

The authors thank the three anonymous reviewers, Cognitive Systems Laboratory members, and Bryan Reimer for their comments and suggestions. The data presented in this article were collected as part of a project sponsored by General Motors under the guidance of Linda Angell. This work was supported by the National Science Foundation through a graduate research fellowship to Shannon C. Roberts.
KEY POINTS

- Tasks with high visual, manual, and cognitive demand, such as searching through long playlists on an MP3 player, lead to long glances away from the road and substantial degradation in driving performance.
- Playlist length is a key design attribute that influences the degree of interference with driving. Long playlists (580 songs) produced the greatest interference with driving, whereas few differences were found between short playlist, medium playlist, radio tuning, and no-task conditions.
- Aftermarket controllers designed to adapt MP3 players to vehicles can undermine rather than enhance performance, as they require more long glances.
- Although drivers hesitated to begin tasks and took shorter glances away from the roadway, they did not successfully modulate their interaction with devices to adapt to roadway demands, as demonstrated by poor driving performance and long glances away from the roadway toward the end of long tasks.

REFERENCES


John D. Lee is a professor of industrial and systems engineering and director of the Cognitive Systems Laboratory at the University of Wisconsin-Madison. He received his PhD in mechanical engineering in 1992 from the University of Illinois at Urbana-Champaign.
Shannon C. Roberts is a PhD student and graduate research assistant in the Industrial and Systems Engineering Department at the University of Wisconsin–Madison. She received her MS in industrial engineering from the University of Wisconsin–Madison in 2011.

Joshua D. Hoffman is an advanced user-interface technology engineer at Deere and Company in Moline, Illinois, and a PhD candidate in industrial engineering at the University of Iowa. He received his MS in industrial engineering from the University of Iowa in 2004.

Linda S. Angell was a technical fellow at General Motors at the time of this work (prior to her retirement in 2008). She is now a visiting research scientist at Virginia Tech Transportation Institute, in its Center for Automotive Safety Research, and a principal at Touchstone Evaluations, an independent testing and research services firm. She received her PhD in experimental psychology from the University of Colorado–Boulder, in 1979.

Date received: February 1, 2011
Date accepted: October 10, 2011