Examining the Efficacy of Training Interventions in Improving Older Driver Performance


Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2012 56: 144
DOI: 10.1177/1071181312561007

The online version of this article can be found at:
http://pro.sagepub.com/content/56/1/144
Examining the Efficacy of Training Interventions in Improving Older Driver Performance

John G. Gaspar¹, Mark B. Neider², Daniel J. Simons¹, Jason S. McCarley³, & Arthur F. Kramer¹

¹Department of Psychology and Beckman Institute, University of Illinois at Urbana-Champaign
²Department of Psychology, University of Central Florida
³Flinders University, South Australia

An increasing number of commercial training products claim to improve older driver performance by training underlying cognitive abilities. However, research examining transfer of such training to driving performance is limited. The current study examined whether 16 hours of training on a commercial training package improved older adults’ performance in a high-fidelity driving simulator. Data showed no differential improvements between the training group and a control group on any driving performance measure following training. The commercial training program did not improve the simulated driving performance of older adults.

INTRODUCTION

Per mile driven, drivers over age 65 are at a heightened risk for accident involvement (NTSA, 2008). Older adults can adjust their driving behavior to compensate for their performance declines, but maintaining the ability to drive safely is critical to prolonged independent living.

To this end, research has examined whether training can improve older driver performance. Effective driver training programs often give drivers practice and feedback on driving tasks, typically using a high-fidelity driving simulator or on-road test vehicle. Romoser and Fisher (2009), for example, examined whether simulator training could improve older drivers’ scanning behavior at intersections in a driving simulator. Their training group received customized feedback and instruction on failures to properly scan critical areas at simulated intersections. A control group received classroom instruction. Following training, the training group had a significantly greater reduction in scanning errors compared to the control group. Similarly, a combination of on-road and classroom training decreased older driver errors compared to a control group (Marattoli et al., 2007; see also Bedard et al., 2008). However, logistical issues limit widespread use of these programs; the protocols require customized, one-on-one feedback, realistic testing environments, and staff to oversee training.

An alternative to training the driving task itself would be to train the cognitive abilities that undergird or enable safe driving performance. Impaired performance on a number of cognitive tasks has been linked to impaired driving performance (see Anstey et al., 2005 for a review). For example, poorer performance on the UFOV, a measure of visual attention, has been linked to higher accident rates for older adults (Clay et al., 2005). And, performance on a flicker change detection task predicts simulated driving performance for older participants (Hoffman et al., 2005). Although these abilities may be amenable to training, improvements generally are confined to the trained task (i.e. training does not transfer; e.g. Ball et al., 2000).

It is unclear whether cognitive training can transfer to driving performance. Cassavaugh and Kramer (2009) found that training on computer based analogs of components of driving (attention, working memory, and manual control) improved older driver performance in a simulator. Roenker and colleagues (2003) trained older adults (screened for UFOV impairment) on either the UFOV or a simulator training program. The only driving measure showing a benefit of UFOV training was ‘dangerous maneuvers’, which was just one of eight composite scores created from subjective ratings. Importantly, the simulator training group showed benefits on a wider range of subjective measures. UFOV-trained older adults also had shorter choice reaction times following training compared to the driver’s-education group, which the authors suggested would translate into faster stopping times. In a follow-up to the ACTIVE trial, Ball and colleagues (2010) compared accident rates for older adults trained on processing speed (via a UFOV task), memory, or reasoning. Overall, accident rates did not differ between groups following training (though the processing speed group had a reduction in at-fault accidents, this was offset by an increase in not-at-fault accidents). These studies, though a promising first step, provide inconclusive support for the efficacy of cognitive training in improving older driver performance.

Though scientific evidence for transfer of cognitive training remains mixed (see also Boot et al., 2008; Boot, Blakely, & Simons, 2010), the “brain” training subset of the video game industry continues to expand. Many of these programs specifically target age-related cognitive decline. CogniFit’s Personal Coach program, for example, claims to provide, “an effective way to guard against the gradual decline in mental sharpness that is a natural but preventable consequence of aging.” (CogniFit, 2010). A number of programs focus specifically on improving older driver performance with computer-based cognitive training. Two of the most prominent packages are Posit Science’s DriveSharp and CogniFit’s DriveFit. These companies make strong assertions about their programs’ efficacy. PositScience, for example, claims their program cuts accident risk by 50% (Posit Science, 2009). This claim originates from the Ball and colleagues (2010) finding that UFOV-trained drivers had lower at-fault accidents after training. As noted above, however, not-at-fault accidents increased for the UFOV group, meaning that UFOV-trained drivers were actually no less likely to be involved in an accident. Posit Science also claims to decrease stopping time while driving, though...
the only evidence comes from Roenker and colleagues (2003) who found faster choice reaction times for the UFOV group, not faster responses in any simulated or on-road driving measure. Increased independent clinical testing of these cognitive training programs is needed.

The present study asked whether a commercial computer-based cognitive training program would improve older driver performance in a driving simulator. Specifically, driving performance, as assessed before and after training using a high-fidelity driving simulator, was compared for a commercial driver training program and a control group that played card games. The training assessment comprised two realistic and demanding driving simulations and included a number of objective performance measures. If the commercial program was effective in improving older driver performance, we expected to see differential improvement following training for the commercial training group compared to the control group on one or more driving performance measures.

**METHOD**

**Participants**

Forty older adults were recruited from the Urbana-Champaign community (mean age = 74.7, S.D. = 4.9). All participants had a valid driver’s license, normal or corrected-to-normal visual acuity, normal color vision, and a score of 27 or higher on the Mini-Mental Status Exam. Participants reported no heart attacks or strokes within the last year and were taking no medications preventing safe driving.

**Apparatus**

The Beckman Institute Driving Simulator at the University of Illinois ([http://isl.beckman.illinois.edu](http://isl.beckman.illinois.edu)) was used to assess driving performance. The simulator consists of a General Motors Saturn Automobile surrounded by eight screens, creating a highly immersive driving environment. Traffic environments and experimental scenarios were developed using HyperDrive™ Authoring Suite. Data were recorded at 60 Hz. Five PC’s with 19-inch displays were used for training.

**Training Programs**

*Commercial Training Program.* Participants in the commercial training group (N = 20) completed the CogniFit Senior Driver program, a commercially available computer package designed to “maintain and extend” the driving ability of older adults. Using scores from an initial assessment, the program creates an individualized training program for each participant to train 14 aspects of cognition (e.g. divided attention, field of view, visual scanning, working memory, etc). Each training session lasted approximately 30 minutes and participants completed two training sessions each time they came to the lab, for 16 total hours of training.

*Control Group.* Participants in the control group (N = 20) played card games on a computer (Hoyle Card Games; Encore Software, Inc., 2008) during 16 one-hour long sessions.

**Driving Assessment**

Participants completed two simulated driving assessments before and after training. Example screenshots from the driving tasks are presented in Figure 1. The *Hazard Response* task was designed to measure responses to unexpected events in the environment. Drivers traversed a straight two-lane urban road with ambient traffic and pedestrians. Participants were instructed to keep their speed as close to 35mph as possible (except when responding to events) and received warning messages if speed fell below 30mph or above 40mph. Eight critical hazard events occurred in each drive. Hazards consisted of events that could occur in everyday driving and were triggered by driving over a pre-determined location in the environment. There were four types of hazards: pedestrians crossing the street, parked cars pulling out, cars turning in front of the participant, and dogs crossing the street. Events were spaced throughout the environment to make the task as realistic as possible and to limit participants’ ability to predict the events; that is, some events occurred in rapid succession while others were separated by large intervals. Two experimental drives were created using the same base map. The order and spacing of events was reversed between the two drives and the order of drives was counterbalanced across participants. The primary performance measure was response time to the hazard events.

The second driving assessment was the *Highway* driving task (Horrey & Simons, 2007). The environment consisted of three lanes of traffic headed in the same direction on a highway. Participants were instructed to merge onto the highway from an off ramp and to maintain a speed of 55mph. Each of the lanes was populated with vehicles that were assigned random velocities within a range that differed between lanes; right (slow) lane 47-52mph, middle lane 52-57mph, left (fast) lane 57-62mph. Vehicles changed lanes at random times and assumed a speed within the new lane range. To create high congestion situations requiring action by the participant, cars ahead of the driver were programmed to occasionally slow down (within the range for the lane) and cars behind the participant were programmed to occasionally speed up (within the lane range). This generated periods of steady-state following behavior where the participant followed behind a lead vehicle in the same lane for an extended period and times when drivers had to pass slower vehicles. The primary measures of performance were steady-state following distance and safety margins when drivers merged to a new lane.
Procedure

Participants completed an initial screening session during which demographic and vision measures were assessed. They then completed pre-training hazard response and highway drives and were randomly assigned to the commercial training group or the control group, where they completed 16 1-hour training sessions. Following training, participants completed hazard and highway driving assessments. The program took an average of 8-9 weeks to complete and participants were paid $10 per session.

RESULTS

Of primary interest was whether the commercial training group showed differential improvement on any driving measure compared to the control group. Analyses were conducted as ANOVAs with training group as a between subjects factor and session (pre/post-training) as a within subjects factor. The reported statistical values include conventional test statistics, and session (pre/post-training) as a within subjects factor. The overall RT’s and RT for each event type are shown in Table 1. For overall RT, there were no main effects of group \((F(1, 28) = 2.81, p = .11, \eta^2_p = .43, \eta^2_{pBC}(H_1|D) = .80)\) or session \((F(1, 28) = 1.52, p = .23, \eta^2_p = .05, \eta^2_{pBC}(H_1|D) = .28)\) and the interaction between group and session was not significant, \(p = .99\).

Table 1. Pre- and Post-Training Driving Results for the Commercial Training and Control Groups

<table>
<thead>
<tr>
<th>Hazard Response</th>
<th>Pre-Training</th>
<th>Post-Training</th>
<th>Pre-Training</th>
<th>Post-Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Response Time (s)</td>
<td>3.92 (.76)</td>
<td>3.47 (.81)</td>
<td>4.05 (.87)</td>
<td>3.75 (.92)</td>
</tr>
<tr>
<td>Pedestrians crossing (s)</td>
<td>2.99 (.70)</td>
<td>2.72 (.70)</td>
<td>3.10 (.76)</td>
<td>2.88 (.78)</td>
</tr>
<tr>
<td>Cars pulling out (s)</td>
<td>2.45 (.92)</td>
<td>2.01 (1.19)</td>
<td>2.33 (1.24)</td>
<td>1.98 (1.27)</td>
</tr>
<tr>
<td>Cars turning in front (s)</td>
<td>4.44 (1.47)</td>
<td>3.97 (1.46)</td>
<td>5.33 (1.07)</td>
<td>4.98 (1.53)</td>
</tr>
<tr>
<td>Dogs crossing (s)</td>
<td>5.41 (1.14)</td>
<td>5.20 (1.05)</td>
<td>5.46 (1.07)</td>
<td>5.14 (1.25)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Highway</th>
<th>Pre-Training</th>
<th>Post-Training</th>
<th>Pre-Training</th>
<th>Post-Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Following Distance (m)</td>
<td>45.57 (8.76)</td>
<td>48.73 (8.56)</td>
<td>48.12 (7.90)</td>
<td>51.85 (11.55)</td>
</tr>
<tr>
<td>Headway Safety Margin (m)</td>
<td>60.0 (13.8)</td>
<td>63.0 (15.1)</td>
<td>59.3 (20.5)</td>
<td>62.6 (18.8)</td>
</tr>
<tr>
<td>Tailway Safety Margin (m)</td>
<td>50.6 (11.8)</td>
<td>48.6 (8.3)</td>
<td>52.7 (15.6)</td>
<td>53.2 (12.2)</td>
</tr>
</tbody>
</table>

Note: Data are displayed as mean (S.D.). Hazard responses indicate time from event trigger until driver response for each event type.

Group differences were examined with ANOVAs with group as a between subjects factor and session (pre/post-training) as a within subjects factor. The effect size measure \(\eta^2_p\), along with \(\eta^2_{pBC}(H_1|D)\), an estimate of the posterior probability of the alternative hypothesis given the observed data (Masson, 2011; Wagenmakers, 2007). Thus, a value of \(\eta^2_{pBC}(H_1|D)\) less than .5 favors the null hypothesis, and a value of greater than .5 favors the alternative.

Hazard Response Driving Assessment

Ten subjects (4 commercial, 6 control) were unable to complete the hazard response drive due to simulator sickness. Data from these subjects were excluded from analyses. Hazard response time (RT) was defined as the time from the triggering of an event until the driver made a brake or steering (outside of 1 standard deviation) response. Only responses in which a collision did not occur were included in the analyses. Overall RT’s and RT for each event type are shown in Table 1. For overall RT, there were no main effects of group \((F(1, 28) = 2.81, p = .11, \eta^2_p = .43, \eta^2_{pBC}(H_1|D) = .80)\) or session \((F(1, 28) = 1.52, p = .23, \eta^2_p = .05, \eta^2_{pBC}(H_1|D) = .28)\) and the interaction between group and session was not significant, \(p = .99\).

The main effects of group \((F(1, 28) = 0.70, p = .41, \eta^2_p = .02, \eta^2_{pBC}(H_1|D) = .20)\) and session \((F(1, 28) = 1.33, p = .26, \eta^2_p = .05, \eta^2_{pBC}(H_1|D) = .28)\) and the group by session interaction \((F(1, 28) = 0.02, p = .90, \eta^2_p = .001, \eta^2_{pBC}(H_1|D) = .16)\) did not reach significance.

Cars turning in front. The main effect of group was significant \((F(1, 28) = 5.34, p = .03, \eta^2_p = .16, \eta^2_{pBC}(H_1|D) = .71)\), with the commercial group responding faster in the pre- and post-training sessions. However, the main effect of session \((F(1, 28) = 1.78, p = .19, \eta^2_p = .06, \eta^2_{pBC}(H_1|D) = .32)\) and the group by session interaction \((F(1, 28) = 0.04, p = .85, \eta^2_p = .001, \eta^2_{pBC}(H_1|D) = .16)\) did not approach significance.

Cars pulling out. The main effect of group \((F(1, 28) = 0.16, p = .67, \eta^2_p = .006, \eta^2_{pBC}(H_1|D) = .17)\) and session \((F(1, 28) = 1.11, p = .30, \eta^2_p = .04, \eta^2_{pBC}(H_1|D) = .25)\) and the group by session interaction \((F(1, 28) = 0.02, p = .90, \eta^2_p = .001, \eta^2_{pBC}(H_1|D) = .16)\) did not reach significance.

Pedestrians. The main effects of group \((F(1, 28) = .70, p = .41, \eta^2_p = .02, \eta^2_{pBC}(H_1|D) = .20)\) and session \((F(1, 28) = 1.33, p = .26, \eta^2_p = .05, \eta^2_{pBC}(H_1|D) = .28)\) and the group by session interaction \((F(1, 28) = 0.02, p = .90, \eta^2_p = .001, \eta^2_{pBC}(H_1|D) = .16)\) did not reach significance.

Cars turning in front. The main effect of group was significant \((F(1, 28) = 5.34, p = .03, \eta^2_p = .16, \eta^2_{pBC}(H_1|D) = .71)\), with the commercial group responding faster in the pre- and post-training sessions. However, the main effect of session \((F(1, 28) = 1.78, p = .19, \eta^2_p = .06, \eta^2_{pBC}(H_1|D) = .32)\) and the group by session interaction \((F(1, 28) = 0.04, p = .85, \eta^2_p = .001, \eta^2_{pBC}(H_1|D) = .16)\) did not approach significance.

Dogs crossing. The main effects of group \((F(1, 28) = 0.001, p = .97, \eta^2_p = .00, \eta^2_{pBC}(H_1|D) = .16)\) and session \((F(1, 28) = 0.49, p = .49, \eta^2_p = .02, \eta^2_{pBC}(H_1|D) = .80)\) and the group by session interaction \((F(1, 28) = 0.02, p = .89, \eta^2_p = .001, \eta^2_{pBC}(H_1|D) = .20)\) were not significant.

The lack of group by session interactions in both the overall and individual task RT’s indicates that commercial training did not differentially reduce RT for any of the event types compared to the control group.
Highway Driving Assessment

Following Behavior. Steady-state following periods were defined as times in which the driver was within 90 meters of the vehicle ahead and the rate of increase or decrease in headway (between the driver’s vehicle and a vehicle ahead) was below 1.5 m/s for at least 5 seconds. This excluded instances where the driver rapidly passed another vehicle without following for an extended period or where the lead vehicle was traveling faster than the driver. Two drivers did not engage in following within 90m of the lead vehicle and were excluded from analyses. The primary measure of performance was average steady-state headway distance. The main effect of group was not significant, $F(1, 36) = 1.49, p = .23, \eta^2_p = .04, p_{BIC}(H_0|D) = .26$. There was a marginally significant effect of session, with larger headway distances in the post-training session for both groups, $F(1, 36) = 3.22, p = .08, \eta^2_p = .08, p_{BIC}(H_0|D) = .44$. Importantly, the lack of an interaction between group and session indicates that commercial training did not result in significant changes to following behavior compared to the control group, $F(1, 36) = 0.02, p = .88, \eta^2_p = .001, p_{BIC}(H_0|D) = .14$. Thus, even though both groups showed some increase in headway in the post-training session, the commercial training group did not differentially increase their headway compared to the control group.

Lane Changing Behavior. The primary measure of performance was the distance from the driver to vehicles ahead and behind in the new lane (Figure 2) when drivers initiated a lane change, defined as the time when half the driver’s vehicle had entered the new lane. Larger safety margins allow the driver more time to respond to unexpected braking (by the vehicle ahead) or speeding (by the vehicle behind). Drivers did not increase their headway ($F(1,38) = 0.79, p = .38, \eta^2_p = .02, p_{BIC}(H_0|D) = .19$) or tailway ($F(1,38) = 0.09, p = .76, \eta^2_p = .002, p_{BIC}(H_0|D) = .14$) safety margins between sessions. Importantly, there were no differences between groups on either headway ($F(1,38) = 0.02, p = .89, \eta^2_p = .001, p_{BIC}(H_0|D) = .14$) or tailway ($F(1,38) = 1.25, p = .27, \eta^2_p = .03, p_{BIC}(H_0|D) = .23$) safety margins and there were no interactions between group and session for headway ($F(1,38) = 0.001, p = .97, \eta^2_p = .00, p_{BIC}(H_0|D) = .14$) or tailway ($F(1,38) = 0.27, p > .61, \eta^2_p = .007, p_{BIC}(H_0|D) = .15$). This indicates that drivers in the commercial training group did not differentially modify their safety margins following training.

DISCUSSION

The present study examined the efficacy of a commercial driver training package. Specifically, older drivers were trained for 16 hours on a commercial package; pre- and post-test performance measures for the trained drivers were compared with those of a control group. Data showed no evidence of differential improvement for the commercial training group compared to the control group on any driving measure. Drivers in the commercial training group did not respond significantly faster in the hazard response task following training compared to the control group, nor did they modify their or safety margins when merging relative to the control group. Both groups showed a marginal increase in steady-state following distance following training, but the increase did not differ between groups.

Research has produced few examples of performance transfer from training on basic cognitive tasks to more complex ones (e.g. Gopher et al., 1994; Roenker et al., 2003). Generally, cognitive training improves trained task performance but does not transfer broadly (e.g. Ball et al., 2002). Our data suggest that training on cognitive tasks within a commercial program is an ineffective means of improving two critical aspects of driving performance. While cognitive abilities such as attentional breadth (e.g. UFOV; Clay et al., 2005) are related to driving performance, driving is a complex task that requires not only visual attention but also manual control of the vehicle, route planning, and knowledge of where potential hazards may occur. Safe drivers must also efficiently allocate attention among these tasks. Cognitive abilities such as the UFOV reflect only a subset of the skills needed for safe driving.

The mismatch in complexity between the training and transfer tasks likely contributes to a lack of transfer. Cognitive training requires that improvements transfer broadly whereas on-road or simulator training programs (e.g. Romoser & Fischer, 2009) allow drivers to acquire skills directly related to driving and to practice them in the testing environment (i.e. near transfer). A promising and more logistically feasible approach may be to train drivers to make judgments about schematic views of driving scenes. Pollatsek and colleagues (2006) trained young, inexperienced drivers using top-down diagrams of dangerous driving scenes on a computer. Trainees marked the areas in each scene where they should scan and where hazards are likely to develop, and then received individualized feedback about their errors and practice remarking the diagrams. A simulator assessment following training indicated that drivers trained on the program were more likely than untrained drivers to scan critical areas. Importantly, the performance benefit was found for new drives that were not part of the training session. Further research is needed to determine whether these are effective for training older drivers.

A key question is why we failed to replicate the pattern of results of Roenker and colleagues (2003), which suggested that UFOV training transfers to improved driving. First, their participants trained specifically on the UFOV paradigm, whereas drivers in the present study trained on the commercial DriveFit program. Though “Field of View” was one of the

![Figure 2. Illustration of lane change safety margin measures for a merge to the left. The black car represents the driver and the white cars represent vehicles ahead and behind in the new lane.](image-url)
cognitive abilities DriveFit claims to train, it remains possible (though perhaps unlikely given the cognitive training and transfer literature) that UFOV-specific training is more efficacious in improving driving. Note, though, that Roenker et al (2003) tested many different outcome measures and found only one that showed any reliable benefit from UFOV training, and that outcome measure was based on an observer’s judgment rather than an objective measurement. The present study examined objective measures of driving performance in two demanding driving simulator assessments. Our analysis allowed for a more fine-grained examination of driving performance. Though some inherent mismatch exists between simulations and on-road driving, using a driving simulator allowed us to put drivers in demanding situations and to measure performance variables which could result in accidents. Our results suggest drivers were no less ‘dangerous’ following training; they did not stop faster or increase headway or safety margin distance.

Another important difference deals with sampling differences. Roenker and colleagues included only older adults screened for UFOV impairment while the present study included a wider range of older adults. Older adults with low UFOV scores represent a subset of the older adult population at increased risk for accidents (see Clay et al., 2005). These drivers may have the most to gain from cognitive training programs. More research is needed to determine the role of individual differences in transfer to driving performance.

Roenker and colleagues (2003) also found faster choice reaction times for the UFOV-trained group, and suggested that this difference translates to faster stopping while driving. We found that drivers who undertook a commercial cognitive training program were no faster in responding to hazard events following training. Additional independent clinical testing is needed to examine whether training does indeed decrease response times in actual or simulated driving tasks. Although additional training might have yielded driving improvements, our data provide no evidence of differential improvement in any of our performance measures after 16 hours of training.

Our results highlight the need for further independent clinical assessment of the claimed benefits of commercial cognitive training programs. Of particular concern is that commercial programs could lead to older driver overconfidence. A testimonial from the DriveSharp website reads, “75-year-old [man] says DriveSharp has improved his peripheral vision [and] made him feel more confident behind the wheel.” Indeed, commercial cognitive training decreases driving cessation and increases confidence (Edwards et al., 2009). However, if these commercial packages fail to improve driving ability, drivers may overestimate their ability and put themselves in overly demanding situations.

In summary, our results provide evidence that a commercial computer-based cognitive training package did not enhance older driver performance. Older adults looking to improve their driving may be better served by seeking out training programs that provide practice in a driving context than by practicing basic cognitive tasks.

ACKNOWLEDGEMENTS

The authors would like to thank the Illinois Simulator Lab for their support and E.M. and V.R. for their help in data collection.

REFERENCES


