Speed-of-Processing and Driving Simulator Training Result in Improved Driving Performance

Daniel L. Roeker and Gayia M. Cissell, Western Kentucky University, Bowling Green, Kentucky, and Karlene K. Ball, Virginia G. Wadley, and Jerri D. Edwards, University of Alabama at Birmingham, Birmingham, Alabama

Useful field of view, a measure of processing speed and spatial attention, can be improved with training. We evaluated the effects of this improvement on older adults' driving performance. Elderly participants in a speed-of-processing training program (N = 48), a traditional driving training program performed in a driving simulator (N = 22), or a low-risk reference group (N = 23). Before training, immediately after training or an equivalent time delay and after an 18-month delay each participant was evaluated in a driving simulator and completed a 14-mile (22.5-km) open-road driving evaluation. Speed-of-processing training, but not simulator training, improved a specific measure of useful field of view (UFOV8), transferred to some simulator measures, and resulted in fewer dangerous maneuvers during the driving evaluation. The simulator-trained group improved on two driving performance measures: turning into the correct lane and proper signal use. Similar effects were not observed in the speed-of-processing training or low-risk reference groups. The persistence of these effects over an 18-month test interval was also evaluated. Actual or potential applications of this research include driver assessment and/or training programs and cognitive intervention programs for older adults.

INTRODUCTION

Some older drivers are prone to crashing, whereas others remain crash free throughout their driving careers. Until fairly recently, little has been known about which behavioral and physiological changes differentiate competent older drivers from those experiencing a decline in driving ability. However, numerous studies in the literature (e.g., Ball, Owsley, Slame, Roeker, & Brun, 1993; Ball et al., 2002; Good et al., 1998; Owsley et al., 1998; Owsley, Ball, Slane, Roeker, & Brun, 1991; Rizzo & Dingus, 1998; Rizzo, Koehler, McGovern, & Dawson, 1997) demonstrate that the driving performance of older drivers is related to a specific functional measure of visual processing speed, the useful field of view test (UFOV8). Furthermore, Ball and his colleagues (e.g., Ball, Ball, et al., 1988; Ball & Owsley, 1991; Edwards et al., 2002) have shown that the size of the useful field of view can be expanded with training. Therefore, the present investigation was designed to determine if speed-of-processing training also produces an improvement in driving performance. Given that 17% (5 million) of the population is expected to be over age 65 by the year 2020 (U.S. Department of Transportation, 1997), the development of techniques to help older adults maintain their mobility and lessen the risk of crash involvement has clear societal benefits.

The useful field of view has been defined as the area from which one can extract visual information in a single glance without eye or head movement (Sanders, 1970). It is measured binocularly and can involve the detection, localization, or identification of targets against more complex visual backgrounds (Verissimo et al., 1983, 1985; Sanders, 1970). The size of the useful field of view is affected by many factors.
including visual sensory function (Owsey, Ball, & Keehn, 1993), a slowing of processing ability, the inability to divide attention, and the inability to ignore distractors (Ball, Roenker, & Brown, 1990). Measures of these last three factors have been incorporated into the UF0V® test administered on the Visual Attention Analyzer (see Ball et al., 1990). The UF0V® test measures the speed at which individuals can process information within a 30° radius visual field under a variety of cognitively demanding conditions.

As noted earlier, recent retrospective (Ball et al., 1993; Goode et al., 1998; Owsey et al., 1998; Owsey et al., 1991) and prospective studies (Ball et al., 2002; Owsey et al., 1998) have shown that older drivers with UF0V® impairments are involved in at-fault crashes more frequently than are drivers without these deficits. Other lines of evidence have supported this relationship. For example, Rizzo and colleagues (Rizzo & Dingus, 1996; Rizzo et al., 1997) have shown a similar relationship between UF0V® impairments and simulator crashes. In another simulator study, Chapparro and his colleagues (Chapparro, Alton, Sifrit, & Geoff, 2001; Sifrit, Chapparro, Groff, & Stumpfhauser, 2001) have shown that individuals with UF0V® impairments encounter more difficulty in detecting peripheral targets and reacting to them in a driving simulator than do unimpaired individuals. Finally, Myers, Ball, Kalina, Roth, and Goode (2000) have shown that there is a linear relationship between the degree of UF0V® impairment and the probability of passing an on-the-road driving evaluation. Thus the relationship between UF0V® performance and driving performance has been well established, and interest has turned to ways to minimize impairments and thereby reduce the risk for crash involvement.

There is abundant evidence that older adults can benefit from perceptual and/or speeded task training (e.g., Kramer, Larish, & Strayer, 1995; Kramer, Larish, Weber, & Bardell, 1999) for example, Kramer and his colleagues have shown that after "variable priority practice," in which the relative attention applied to two tasks is varied in a non-systematic fashion across practice sessions, older adults' performance on subsequent attention-switching tasks approaches that of younger participants. This ability to rapidly switch attention between multiple tasks clearly seems relevant for driving. Furthermore, the transfer of perceptual/cognitive training to simulated and actual driving performance is not unprecedented! Sifrit et al. (2001) found that a form of speed-of-processing training led to greater situational awareness (number of hazards detected) in a simulated driving task. Sivak et al. (1984) found that training individuals on a series of elementary perceptual tasks (visual scanning, figure-ground discrimination, etc.) led to improved open-road driving performance in a small sample of individuals with brain damage. Although it is unclear which aspect of the training was responsible for subsequent changes in driving behavior, their results demonstrate that training of fundamental perceptual/cognitive skills can improve driving performance.

In the process of exploring the factors that affected UF0V® test performance, Ball, Beard, Roenker, Miller, and Griggs (1988) developed a speed-of-processing training program that has been shown to expand the size of the useful field of view. The benefits of this training protocol have been shown to generalize to untrained stimuli over shorter stimulus durations (Ball, Ball, et al., 1988; Ball, Beard, Roenker, Miller, & Ball, 1988; Ball et al., 1991; Sekuler & Ball, 1986). That there is clear evidence that certain types of training can produce cognitive changes that have the potential to improve driving performance. However, the detection of such benefits is problematic because of the difficulties encountered in trying to define and measure driving competence. There have been three major methods for assessing driving skills: crash history, driving simulator performance, and on-the-road driving performance. Each of these methods has advantages and disadvantages. Crashes are a relatively insensitive measure of driving performance because they are rare events and often go unreported by either the individual (Owsey et al., 1991) or the reporting agency (i.e., some states do not record crashes that occur on private property). Driving simulators offer experimental control over the driving experience, but they often lack high fidelity with actual driving and may result in simulator sickness, particularly
for older adults. Road tests are frequently con-
sidered to be the best measure of driving com-
petence, but they also have their drawbacks,
particularly with respect to inconsistencies in
the administration and scoring of the results.
For example, road tests vary in traffic density
(Herbert, 1963; Summala, Nieminen, & Punto,
1996), number of evaluators (Odenheimer et
al., 1994; Sivak et al., 1984), sampling period
(Hunt, Morris, Edwards, & Wilson, 1995; Sivak
et al., 1984), and rating systems (Hunt et al.,
1993).

The present study used an on-the-road driv-
ing evaluation that attempts to minimize the
methodological limitations often associated
with this outcome measure. In particular, a com-
plex rating system involving 455 operationally
defined behaviors was developed. Second, multi-
ple evaluators were used and their reliability
was assessed. Finally, the route was repeated
twice to provide an opportunity for a range of
traffic conditions to occur. The present study
attempts to detect changes in on-the-road driv-
ing behavior induced by a form of cognitive
practice, speed-of-processing training.

METHOD

Participants

The participants were 456 licensed drivers
(187 men, 269 women; 428 Caucasians, 28
African Americans) with an average age of 69
years (range 48–94 years) who were screened
for participation in the training study. The par-
ticipants were residents of Warren County,
Kentucky, and the surrounding area and were
recruited using a variety of methods (letters sent
by the Kentucky Transportation Cabinet to
crash-involved drivers age 55 and older, tele-
phone calls, and talks to community groups).

In order to identify older drivers who would
potentially benefit from speed-of-processing
training, screening measures of visual acuity,
contrast sensitivity, and UFOV® were conduct-
ed. Visual acuity (Lighthouse Distance Visual
Acuity Test, second edition, Ferris, Kassoff,
Brennick, & Bailey, 1982) and contrast sensi-
tivity (Pelli, Robson, & Wilkins, 1988) were
assessed binocularly using the participant's cor-
rective lenses. Inclusion criteria for the training
study were a Bailey-Lovie acuity score (Bailey
& Lovie, 1980) of 50 logMAR or better and a
log contrast sensitivity of 1.3 or better (Bailey
& Lovie, 1980); these values have been shown by
Owsley et al. (1993) to be the minimal visual
skills necessary to perform the UFOV® task.
Participants who failed the visual screening (n =
2) were referred to a vision specialist and exclud-
ed from participation.

Study inclusion criteria also included UFOV®
reduction. Inclusion in either the speed-of-
processing or simulator training group required
a minimum of 50% total reduction on the
UFOV® measure. Prior research (Ball et al.,
1993; Owsley et al., 1991) has shown that a cut
point of 40% reduction provides maximum
sensitivity and specificity in predicting crash
involvement from state records. Crash involve-
ment, however, is a relatively insensitive measure
of driving performance, and given that the pri-
mary measure of driving performance in the
present investigation was on-the-road driving,
the more lenient cut point (30% reduction) was
taken.

The goal of recruitment into the training
study was to include a minimum of 50 individu-
ers exhibiting UFOV® decline in the "high-risk"
speed-of-processing training group (the training
of interest), 25 individuals exhibiting similar
decline in a driving simulator training control

group (traditional driver training), and 25 indi-
viduals who did not exhibit UFOV® decline in
a "low-risk" reference group. Although our
prior research has clearly shown that speed-of-
processing training results in a robust increase
in the size of the useful field of view, we were
concerned that the transfer of any such bene-
fits to a skill such as driving might be difficult
to detect. Hence we doubled the sample size
for the speed-of-processing training group,
which resulted in sufficient power (e.g., .80) to
detect effect sizes as small as .35. Additionally,
our primary interest was in ascertaining any such
benefits relative to a sham training control
group. The traditional driver training program
employed in the current research utilized a non-
interactive driving simulator, and we expected
that any benefits derived from such training
would be both minimal and short lived, given
that our participants were experienced drivers.
Finally, the low-risk reference group was includ-
ed to serve as a second control group against
which the baseline driving performance of the two high-risk groups could be compared.

The degree of UFQG decline was assessed on the Visual Attention Analyzer. Before UFQG testing began, all participants demonstrated the necessary sensory capacity to correctly locate at least 75% of the peripheral targets.

Training Participants:
In the training study were 104 participants (58 men, 66 women, 100 Caucasians, 4 African Americans) with a mean age of 71 years (range 55-86). Of the 496 individuals screened, 3 failed to complete the protocol and 2 were excluded for visual impairment. From the 276 participants who had a UFQG reduction of less than 50% (i.e., UFQG < intact), 21 were randomly selected for the low-risk reference group. Of the 75 participants with a UFQG reduction of 50% or greater, 34 declined to participate, and 77 of the remaining 141 were selected for participation so as to create a range of UFQG performance. Within this group, 26 were assigned to the simulator training control group and 51 were assigned to the speed-of-processing training group. All participants were paid for their participation.

Procedure:
Driving simulator measures. Participants' simple (SRT) and choice reaction time (CRT) were assessed in a driving simulator (Model L-225, Dorcon Precision Systems, Inc., Binghamton, NY) consisting of a 35-mm projection system and five driving consoles (steering wheel, brake and accelerator pedals, and an instrumented dashboard).

For SRT, the participants were instructed to watch a light arrangement containing 16 colored lights (1 x 1 cm) located on the top panel of the driver's unit. Color-matched pairs of these lights blinked on for 1 s and off for 0.5 s in a random order. The driver was instructed to brake as quickly as possible when the two red lights (simulated brake lights) were illuminated. After at least 3 practice trials, 15 experimental trials were performed. The duration for each trial was the elapsed time between the onset of the brake lights and the release of the accelerator pedal.

For CRT, a narrated film from the Dorcon film library was viewed at a distance of 5.8 meters. The stimuli (44 x 46 cm, or 4.3° x 4.5°) were road signs (pedestrian, bicycle, right and left turn arrows) with and without a red slash through them. Participants were instructed to react only to signs without a red slash, by braking for a bicycle or pedestrian sign or by turning the steering wheel in the appropriate direction for right or left turn arrows. A value of 5.1 s was assigned for trials in which the action signal was not detected because this was the length of time the display remained visible. The number of stimuli ranged from three to six signs, and within a trial this number was held constant, although the positions of the signs changed throughout the frames. Two identical sets of 10 trials were completed.

On-the-road driving evaluation. The driving evaluation course was designed based on a review of the accident literature, traditional driver's tests, and simulator-to-drive evaluations. The course consisted of two loops of a 7-mile (11.5 km) urban/suburban route. To aid rats in completing their evaluations, maps in the raters' manuals included landmarks at each intersection that were used to define the start and end points for evaluating specific driving behaviors. Finally, analysis of the older driver literature (e.g., Odenheimer et al., 1994) indicated that certain maneuvers, such as left turns across traffic, are especially difficult for older drivers. These locations in the drive were thus considered to constitute places of potential danger. For each of these locations, an item was included to code the extent to which the driver's behavior constituted a dangerous maneuver. A dangerous maneuver was defined as one in which either the driving instructor had to take control of the car or other vehicles had to alter their courses in order to avoid a collision.

The raters' evaluation manual contained 455 items. In order to minimize the demands of the evaluation task, the three raters practiced evaluating drivers on each course segment until all scoring criteria could consistently be applied by all raters. This training was effective, as evidenced by the high (r > 0.92) interrater
 reliability achieved during the road test evaluation (see Results section). The evaluators were required to note only inappropriate driving behaviors. A pilot study was performed on the road test using 4 elderly volunteers to ensure that the course was not too difficult.

The open-road driving evaluation was performed in a car modified with a passenger-side brake for the driving instructor use. Before beginning the evaluation, the driver was familiarized with the car. The driving instructor reformed the driver that he would provide route instructions and that the driver should operate the car as he or she normally would. The driving instructor followed a script in order to give standardized directions. The driver was then directed through a 1-mile (+1.6 km) warm-up route. The subsequent driving evaluation consisted of two loops of a 7-mile (+11.3 km) urban/suburban course. The evaluation was performed during the day (between 7:30 a.m. and 5:00 p.m.) and required 50 to 60 min to complete.

Two of the three independent backseat evaluators rated each driver. All possible combinations of the three evaluators were used. For the pretraining drive, the driving instructor and backseat evaluators were blind to the training condition assigned for all participants. On the posttraining drive, the driving instructor was blind to whether participants were in the speed-of-processing training group or the low-risk reference group. In addition, for approximately half of the drivers, one of the backseat evaluators was blind to the condition to which the participant had been assigned. By comparing the responses of each pair of evaluators, it was possible to guard against any potential bias during the driving evaluation. An analysis of interrater reliability showed that the reliability was equally high when both evaluators were blind (r = .92), one evaluator was blind (r = .93), or neither evaluator was blind (r = .92) to the drivers’ training condition.

All driving behaviors except one were rated on a 4 (very unsafe) to 2 (safe or appropriate) scale. Scale ratings were operationalized for each behavior at each location on the course. Behaviors included maintaining lane position, activating signals, stopping smoothly, searching (additional mirrors were placed in the car to aid the raters in detecting eye movement), selecting gaps, accelerating and decelerating smoothly, turning, maintaining speed, maintaining position in traffic, and dangerous maneuvers. The one behavior rated on a 4-point (from 1 to 4) scale was stopping at stop signs, with 1 defined as running the stop. On a rolling through the stop, 2 as inappropriate stop position, and 3 an appropriate stop.

After the drive, the raters also provided a global rating of the drive (minimum interrater reliability = .84). This global rating ranged from 1 (driver absent/very unsafe) to 6 (very competent driver).

Training. The low-risk reference participants did not receive any training but participated in the same pre- and posttest assessment sessions as did the two high-risk groups. These assessments were separated by an equivalent amount of time (2 weeks) as those of the training groups. As previously described, participants selected for the low-risk reference group were individuals without GPAQ® impairment. This group provided age-matched baseline performance against which we could evaluate the performance of the training groups.

Speed-of-processing training was evaluated relative to a simulator-based training program designed to control for social contact and instructional aspects of training. The simulator training program was representative of a current method of driver retraining and focused on teaching specific driving behaviors. Alternatively, in the speed-of-processing training program, participants did not practice any specific behaviors related to on-the-road vehicle operation, instead, they practiced a fundamental processing skill.

Simulator training control group. A certified driving instructor conducted two educational sessions with groups of 3 to 4 participants each. The first 2 hr session consisted of a review of the general rules of the road and instruction about and simulated practice of safe driving and crash prevention behaviors. Each trainee individually practiced with a driving simulator (a 2300 driving simulator with software demonstrating techniques for crash avoidance, managing intersections, and scanning. Simulator instruction provided participants with skills that were directly applicable to the road test and thus had high face validity for
the participants. The second 2 hr session contin-
ued with the simulation instruction and ended
with a 1 hr in-car demonstration by the driving
instructor of many of the described skills.

**Speed-of-processing training.** Speed-of-
processing training, administered on a touch-
screen computer, was individualized. A detailed
description of training protocols has been pro-
vided elsewhere (e.g., BaI, Beard, Roeneker, Mil-
ner, & Grigg, 1988; BeI et al., 1991; Edwards et
al., 2002), and only a brief summary is provided
here. Based on an initial UFOM® screening, the
threshold for the first UFOM® subtest (stimulus
identification) was determined. If this threshold
was greater than 50 ms, training was imple-
mented by having participants practice the task
at progressively faster presentation speeds until
the participant's threshold was lowered to
17 ms.

The second level of training required the
participant to perform the centrally located
identification task and also to locate a peripheral
target presented as up to 30° of eccentricity. If
the participant was unable to perform these two
tasks at the shortest exposure duration (40 ms),
he or she was further trained, first by practicing
the identification task in conjunction with locat-
ing peripheral targets near the central fixation
box (i.e., 10° of eccentricity) at a presentation
duration slightly faster than threshold. As this
task was mastered, the targets were moved fur-
ther into the periphery (e.g., 20°, 30°). When the
participant could correctly locate the peripheral
target at 30° eccentricity with approximately
75% accuracy, the process was repeated with
near targets (e.g., 10°) and a presentation dura-
tion 40 ms faster than the previous training
speed. This entire process was repeated at pro-
gressively faster presentation speeds until the
participant could perform both the identifica-
tion task and the peripheral localization task
(at 30°) with approximately 75% accuracy at the
fastest presentation speed (80 ms).

The third level of training required the par-
ticipant to perform the identification task and
locate the peripheral target embedded among
distractors. If the participant was unable to per-
form these two tasks at an exposure duration
of 120 ms, training was provided in the same man-
ner as described earlier. Practice continued until
75% correct performance was achieved at an
exposure duration of 120 ms with peripheral
targets at 30°. Thus all participants were trained
to criteria. The average number of training tri-
als completed was 1040 (SD = 504, range
508–3104). The average training time was 4.5
hr. Unlike the simulator training, the speed-of-
processing training had little face validity for
the participants, who routinely questioned the
relationship of such training to driving perfor-
mance.

**Posttraining assessment.** Posttraining mea-
sures were obtained once again in the driving
simulator, on the road, and on UFOM®. For all
groups a 2-week period, on average, elapsed
between the pre- and posttraining assessments.
Following the posttraining driving evaluation,
the participants were offered at individual road
test report prepared by the driving instructor.
Participants were invited to return for follow-up
testing 18 (±1) months after the posttraining eval-
uation, and 60 did return (speed of process-
ing ρ = 28 simulator α = 14, low-risk reference
α = 18). The 18-month assessment procedure
was identical to the pre- and posttraining as-
sessment procedures. A comparison of those
who chose to return and those who did not
revealed no baseline differences in age, speed-
of-processing impairment, reaction time, or
performance on any of the driving composite.

**RESULTS**

Of the 36 participants in the simulator train-
ing group, 22 (mean age = 72.35 years, SD =
5.36, range = 63–81) completed the baseline
and immediate posttest protocol (2 dropped out
because of simulator sickness and 1 because
of equipment failure) and 14 returned for the
18-month follow-up. Of the 57 participants in
the speed-of-processing training group, 48 (mean
age = 72.08 years, SD = 6.82, range = 59–86)
completed the training protocol (2 dropped out
because of illness and 1 had discontinued driv-
ing) and 28 returned for the 18-month follow-
up. Of the 27 participants in the reference
group, 25 (mean age = 69 years, SD = 6.85,
range = 55–80) completed both baseline and
immediate posttesting (2 chose to discontinue
participation), with 14 returning for follow-
up testing. Missing values for the 18-month
follow-up (approximately 17%) were imputed
by inserting, for that individual, the mean immediate posttest value for the respective group for that variable.

In all analyses, gender was initially included as a variable. However, in all cases there was neither a main effect for gender nor any interactions involving gender. Thus subsequent analyses were collapsed across gender.

Analysis Plan

For each of the outcome variables (UFOW®, SRT, CRT, and driving performance) a 3 (groups) x 3 (time) mixed analysis of variance was performed. In addition, in cases where the Group x Time interaction was significant, simple effects analyses and planned comparisons were employed to examine differences between groups at each measurement point. Additionally, Bonferroni’s test (.05) was used to examine changes within groups over time. Recall that the design included a low-risk, unimpaired reference group whose purpose was to serve as a reference point against which to evaluate relative changes in the two impaired groups of interest. Therefore, the guiding principle in the simple effects analyses was to compare the performance of the two high-risk training groups with that of the low-risk reference group.

Before proceeding to these analyses it was necessary to reduce the abundant data from the driving evaluation into a more manageable form. Given the high interrater reliability (.93, SD = .05; .94, SD = .05; and .92, SD = .05 for the pretraining, posttraining, and 18-month follow-up drives, respectively) and the large number of behaviors rated, the 455 items were grouped into 13 composites based on the behaviors being rated. For example, all items that required the evaluator to determine the vehicle's position relative to other traffic were grouped together into a single composite, regardless of where those ratings were made during the course of the drive.

The resulting 13 composites were (a) acceleration: smoothness in use of the accelerator pedal; (b) gap selection: safely merging into or crossing traffic flow; (c) position in traffic: position relative to surrounding traffic while moving; (d) search: eye and head movements at intersections; (e) signals: proper and timely use of turn signals; (f) speed: vehicle speed control relative to posted speed limits; (g) stop position: vehicle position when required to stop at a traffic control device; (h) deceleration: smoothness in vehicle deceleration; (i) tracking: position of vehicle in proper lane; (j) turning: position of vehicle when turning; (k) right of way: yielding to traffic at four-way stops; (l) changing lanes: changing lanes on a multiple-lane road; and (m) dangerous maneuvers: the degree of danger felt by the raters at 17 potentially dangerous locations during the road test. These 17 locations involved high-traffic roadways and were composed of 6 unprotected left turns across traffic, 9 left turn entries to a high-traffic road, and 2 opportunities for inappropriate stopping in traffic to make a right turn.

Unfortunately we were forced to drop five of the composites from the analysis. Performance on four of the composites (gap selection, acceleration, deceleration, and right of way) suffered from ceiling effects. Performance was perfect on approximately 97.5% of the observer opportunities, and mean performance never dropped below 1.94 on a 2-point scale. The search composite was dropped for lack of sufficient data. Although additional mirrors were placed in the vehicle to aid in the detection of eye movements, certain situations (e.g., driver wearing sunglasses, movement of the sun visor) made this assessment difficult. These events frequently took place in middrive, making subsequent evaluation impossible.

For each of the remaining eight composites, scores were generated by averaging both raters' scores (0 = unsafe, 1 = somewhat unsafe, 2 = safe) on all of the items making up that composite. However, the dangerous maneuvers composite was transformed into the actual number of dangerous maneuvers for ease of presentation.

UFOW® test. Means and standard deviations for UFOW®, SRT, and CRT at each of the three testing points are presented in Table 1. An analysis of the UFOW® data (see Figure 1) revealed significant effects for group, \( F(2, 92) = 19.32, p < .001, \) MSE = 118.97, \( \eta^2 = .36 \). time, \( F(2, 184) = 54.59, p < .001, \) MSE = 157.74, \( \eta^2 = .37 \); and a Group x Time interaction, \( F(2, 184) = 40.17, p < .001, \) MSE = 37.74, \( \eta^2 = .47 \). Simple effects analysis of the data revealed that at baseline, as assigned, the low-risk reference group
had significantly smaller UFOV® reduction than did the two training groups. At immediate posttest, however, the speed-of-processing-trained group’s performance was equal to that of the low-risk reference group, whereas the simulator-trained group’s performance remained high. This pattern was still present at the 18-month follow-up. Thus the speed-of-processing training was effective in decreasing the degree of useful field of view reduction, and this training effect persisted at the 18-month follow-up.

SPT. For the simple RT measure there were no effects of group or time, nor was there an interaction (all Fs < 1.49, ps > .05), indicating that this measure was insensitive to group differences or training effects.

CRT. The choice RT data are presented in the top panel of Figure 2. An analysis of these data revealed significant effects for groups, F(2, 92) = 6.02, p < .003, MSE = 0.189, ε² = .12; time, F(2, 184) = 34.23, p < .001, MSE = 0.025, ε² = .27; and a Group × Time interaction, F(2, 184) = 4.35, p < .002, MSE = 0.025, ε² = .08. Simple effects revealed that the low-risk reference group had, as expected, faster CRT at each of the three testing points. Post hoc analyses revealed that the simulator-trained group’s CRT remained relatively stable over time, whereas the speed-of-processing-trained group showed an initial drop at immediate posttest, which was maintained at the 18-month follow-up.

### Table 1: UFOV®, SRT, and CRT Measures: Mean (SD) Baseline, Immediate Posttraining, and 18-Month Posttraining Scores

<table>
<thead>
<tr>
<th>Reference, n = 25</th>
<th>Simulator, n = 22</th>
<th>SOP, n = 48</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Base</td>
<td>Post</td>
</tr>
<tr>
<td>UFOV®</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23.40</td>
<td>27.88</td>
</tr>
<tr>
<td></td>
<td>(4.73)</td>
<td>(7.08)</td>
</tr>
<tr>
<td>Short RT (s)</td>
<td>0.865</td>
<td>0.861</td>
</tr>
<tr>
<td></td>
<td>(0.168)</td>
<td>(0.163)</td>
</tr>
<tr>
<td>Choice RT (s)</td>
<td>1.429</td>
<td>1.752</td>
</tr>
<tr>
<td></td>
<td>(0.223)</td>
<td>(0.239)</td>
</tr>
</tbody>
</table>

* SOP = speed of processing.
On the Road Driving Evaluation

Global rating. Means and standard deviations for the eight driving composites and the global driving rating are presented in Table 2. There was an overall improvement in global rating across testing sessions, \( F(2, 184) = 9.85, p < .001, \) \( \text{MSE} = 0.32, \text{e}^{-2} = 10, \) and a Group \times Time interaction, \( F(4, 184) = 3.26, p = .013, \) \( \text{MSE} = 0.32, \text{e}^{-2} = .07. \) Simple effects analyses revealed that at baseline, as expected, the low-risk reference group's global rating was higher than that of the simulator and speed-of-processing groups. These differences had disappeared at the immediate posttest. However, at the 18-month follow-up the simulator-trained group's global rating was significantly lower than that of the low-risk reference group. Post hoc tests revealed that ratings for the low-risk reference group remained high and did not
| TABLE 2: Global Rating and Driving Evaluation Composites: Mean (SD) Baseline, Immediate Posttraining, and 18-Month Posttraining Scores |
|---------------------------------|-----------------|-----------------|-----------------|
|                                | Reference, n = 25 | Baseline, Post 18 Mo | Simulator, n = 22 | Baseline, Post 18 Mo | SOP*, n = 48 | Baseline, Post 18 Mo |
| Global rating*                 | 4.41 (1.26) | 4.54 (1.16) | 4.68 (1.20) | 3.61 (1.45) | 3.40 (1.53) | 3.63 (1.55) | 4.04 (1.18) | 4.39 (1.19) | 4.25 (1.22) |
| Driving Composites*             | 1.01 (0.83) | 1.03 (1.34) | 1.01 (0.69) | 0.78 (0.14) | 0.84 (0.45) | 0.76 (0.96) | 0.78 (0.26) | 0.82 (0.26) | 0.88 (0.27) |
| Dangerous maneuvers*            | 0.65 (1.15) | 0.75 (1.09) | 1.14 (1.73) | 0.70 (0.14) | 0.74 (0.18) | 0.68 (0.26) | 0.70 (0.26) | 0.74 (0.26) | 0.78 (0.27) |
| Signals                         | 1.83 (0.14) | 1.75 (0.26) | 1.84 (0.20) | 1.74 (0.28) | 1.84 (0.28) | 1.77 (0.28) | 1.74 (0.28) | 1.76 (0.28) | 1.78 (0.28) |
| Turning                         | 1.51 (0.17) | 1.56 (0.18) | 1.57 (0.20) | 1.39 (0.25) | 1.57 (0.28) | 1.53 (0.28) | 1.43 (0.24) | 1.44 (0.24) | 1.55 (0.20) |
| Changing lanes                  | 1.94 (0.28) | 1.80 (0.56) | 1.97 (0.42) | 1.89 (0.44) | 1.89 (0.48) | 1.68 (0.58) | 1.97 (0.12) | 1.94 (0.12) | 1.92 (0.23) |
| Position in traffic             | 1.87 (0.26) | 1.96 (0.16) | 2.00 (0.00) | 1.88 (0.25) | 1.98 (0.11) | 1.98 (0.07) | 1.91 (0.23) | 1.99 (0.05) | 1.88 (0.11) |
| Stop position                   | 1.92 (0.10) | 1.91 (0.10) | 1.91 (0.08) | 1.85 (0.15) | 1.90 (0.12) | 1.86 (0.16) | 1.90 (0.08) | 1.92 (0.08) | 1.90 (0.08) |
| Tracking                        | 1.94 (0.07) | 1.96 (0.05) | 1.95 (0.08) | 1.88 (0.18) | 1.91 (0.18) | 1.85 (0.19) | 1.91 (0.09) | 1.92 (0.09) | 1.91 (0.11) |
| Speed                           | 1.89 (0.18) | 1.88 (0.15) | 1.91 (0.13) | 1.86 (0.15) | 1.81 (0.25) | 1.80 (0.24) | 1.85 (0.15) | 1.88 (0.15) | 1.88 (0.14) |

* SOP = speed of processing; ** Global rating: Scores are averaged across two raters who rate on a 1-6 scale in which 1 = abnormally slow and 6 = competent driver; Driving composites: Composite scores are averaged across two raters on all of the items within the composite. Scores for all but the dangerous maneuvers composite range from 0 to 2 (0 = unsafe, 1 = somewhat unsafe, 2 = safe); The dangerous maneuvers composite is the mean number of dangerous maneuvers observed; thus lower scores reflect better performance.

change across testing sessions, whereas the simulator-trained group was rated poorer at baseline, approached the reference group's overall level of performance at the immediate posttest, and returned to baseline levels at the 18-month follow-up. There was a general, but not significant, increase in global ratings of the speed-of-processing group over time.

Driving composites. The eight driving composites provided an opportunity to determine more precisely the nature of changes in driving performance over time. Five of the eight composites yielded significant Group × Time interactions. These effects are examined in the following paragraphs.

For the dangerous maneuvers interaction, F(4, 184) = 2.89, p < .024, MSE = 0.667, eta2 = .06. The data for the dangerous maneuvers composite is presented in the bottom panel of Figure 2. As expected, the low-risk reference group had fewer dangerous maneuvers during the baseline drive than did either of the high-risk training groups. At immediate posttest, the number of dangerous maneuvers for both training groups dropped and was not significantly different from that of the low-risk reference group. However, at the 18-month follow-up, both the low-risk reference and the simulator-trained groups demonstrated significantly more dangerous maneuvers than did the speed-of-processing trained group.

For the signals interaction, F(4, 184) = 7.22, p < .001, MSE = 0.008, eta2 = .14. The top panel of Figure 3 presents the signal composite data. Simple effects analyses revealed that the low-risk reference group was more consistent in their use of turn signals at baseline than were the two training groups. However, at immediate posttest the low-risk reference group's use of turn signals declined, whereas that of the simulator-trained group improved. This latter effect is not surprising, given that some of the simulator training was directly devoted to proper signaling. These changes did not persist, both groups returned to baseline levels at 18 months. At each assessment from
baseline to 18 months the speed-of-processing-trained group maintained the same level of signal use.

For the turning interaction, $F(4, 184) = 4.30, p < .002, MSE = 0.012, \eta^2 = .09$. The turning composite data are presented in the bottom panel of Figure 3. The low-risk reference group was better than either of the two high-risk groups at turning into the proper lane at baseline. The simulator-trained group showed improvement at immediate posttest, whereas the speed-of-processing group did not, relative to the low-risk reference group. Again, because proper turning was a focal point of the simulator training protocol, this effect is not surprising. The three groups did not differ at the 18-month follow-up.

For the changing lanes interaction, $F(4, 184) = 2.92, p < .023, MSE = 0.073, \eta^2 = .06$. The three groups did not differ at baseline on the ability to correctly change lanes in traffic. However, at immediate posttest there was a
DISCUSSION

The data from the present study demonstrate that both speed-of-processing training and driving simulator training can enhance the driving performance of at-risk older adults but that some of these gains may disappear over time in the absence of any form of "booster" training. Before discussing these results in detail, it is important to examine the methodological improvements of the present study as a context in which to evaluate the current findings.

In designing this on-the-road driving study, an attempt was made to include the in-depths and remove the weaknesses of previous driving studies (de Gier et al., 1981; Hart, Nelenmans, & Bergman, 1981; Hunt, et al., 1993; Odenheimer et al., 1994; Strak et al., 1984). First, in order to get the most realistic measure of driving skills and habits, each participant drove a 7-mile (~11.3-km) route twice. In repeating the route, the driver was allowed to become more comfortable with the evaluation process and therefore was more likely to revert to everyday behavior. Second, instead of using a limited 2-point scale, fall scale to rate behavior, we used a 3-point scale (very unsafe, somewhat unsafe, safe); which allowed for finer gradations in judgment corresponding to the wider range of real-world driving behaviors. Third, the driving behaviors were evaluated by two raters rather than a single individual (e.g., de Gier et al., 1987; Hunt et al., 1993). Very high interrater reliability (.95+), relative to those found by others (e.g., 74-84, Odenheimer et al., 1994), were achieved. Whenever possible, at least one of the raters was blind to the treatment condition of the participant. Ratings by blinded and unblinded evaluators were highly correlated, indicating an objective analysis of driving behavior.

Fourth, a large number of behavioral ratings (475) were obtained in the current study, as compared with previous studies (de Gier et al., 1981; Hunt et al., 1993; Odenheimer et al., 1994), again enabling a fuller range of driving behaviors than generally has been assessed in prior research. These item ratings were grouped into robust and stable categories, each of which was made up of several samples of the driving behavior rather than a single instance. Finally, the evaluations were not limited to certain times...
of the day or to low-traffic conditions, and hence they encompassed a wide range of driving conditions. All these steps were undertaken in order to get the most accurate and least biased measures of driving skills as possible. Clearly, on-the-road driving studies inherently lack control over many circumstances, especially traffic conditions and the behavior of other drivers.

Finally, it is interesting to note that despite our best efforts to measure driving behavior under demanding situations, the performance of the participants in some areas was nearly perfect. Apparently the older adults in our sample, even those in the high-risk groups, had little difficulty waiting for an appropriate opening in traffic (gap selection), waiting for their turn at stop signs (yielding the right of way), or smoothly accelerating and decelerating an unfamiliar vehicle. These findings may be surprising in light of the fact that these tasks are not ones that require rapid decision making or rapid processing of quickly changing information. In fact, most of our participants, even the high-risk ones, drove extremely cautiously, requiring an average of 1 hr to drive 14 miles (22.5 km). As further evidence of this cautiousness, performance on most of the other driving composites was quite high as well. Fortunately, driving failures (e.g., crashes or dangerous maneuvers) are relatively infrequent events, and weaknesses in driving skills may be evident only when the driver is placed in challenging situations that require them to rapidly evaluate and respond. These are precisely the kind of situations that older drivers report avoiding (Odenheimer et al., 1994).

In light of this study's improvements in assessing driving skills, as well as the limitations imposed by the driving environment, we now turn to consideration of the driving results. Although both the speed-of-processing and simulator groups were judged by the independent raters to have improved their driving skills, the nature of improvement for the two types of training varied. The simulator group did not improve on speeded reactions but did improve on a few of the specific driving maneuver skills on which they were trained (turning into the correct lane and signaling), whereas the speed-of-processing and the low-risk reference groups did not. Numerous participants indicated during simulator training that they intended to implement the skills being instructed. The data show, at least in those cases, that they were successful in doing so, although these newly acquired skills had largely disappeared by the time of the 18-month follow-up.

However, the speed-of-processing-trained group improved on untrained tasks that relied on visual attention and higher-order processing speed. The speed-of-processing-trained group improved on the choice reaction time task in the driving simulator, which involved scanning a visual scene, detecting changes in stimuli, and quickly reacting to those changes. For a vehicle moving at 55 miles/hr (88.5 km/hr), this improvement of 277 ms translates into a stopping distance 22 feet (6.7 m) shorter. The speed-of-processing-trained group also made fewer dangerous maneuvers during the postdriving tests than on the baseline drive. The dangerous maneuvers composite consisted of items that primarily tapped visual detection and judgment abilities: in visually cluttered and cognitively demanding high-risk driving situations, such as reading intersecting traffic control devices and making gap selection in order to make turns across oncoming traffic. These same behaviors (or rather the lack of them) are often cited as causes of crashes (Campbell, 1966; Kline, 1986).

Many participants in this training group questioned the link between the training they received and actual driving, and thus they did not expect to improve their driving skills. Taken together, these data indicate that the benefits of training were localized to logically compatible behaviors and that these effects were present over and above any general training effects. The benefits of speed-of-processing training, for the most part, were still present 18 months after training. Furthermore, it is important to note that many, but not all, of the individuals retained the benefits of speed-of-processing training, suggesting that a system of "booster" training may be necessary for some individuals. These results expand on those of Sivak et al. (1984), who measured eight categories of behavior: top acceptance, limit line, observations-turning, observations-straight, path-turning, path-straight, speed-turning, speed-straight. Using a composite
of these measures, they reported an improvement in driving skills after perceptual training. Only composite results were reported, so it is not clear whether participants improved in all categories of behavior or in selected behaviors. The present study measured similar behaviors (gap selection, stop position, right of way, searching, turning, tracking, speed, and acceleration) as well as other behaviors (signals, position in traffic, changing lanes, deceleration) but analyzed for changes within each category of behavior. These analyses did not reveal categorical changes for the speed-of-processing-trained group, but an improvement was noted for the dangerous maneuvers composite, which included behaviors similar to those measured by Stav et al. Thus, our results, like those of Stav et al., support the argument that functional benefits are derived from improving the speed with which one processes complex visual information, given that perceptual skills underlie many daily activities (e.g., driving, walking, general mobility in the environment). These findings are also consistent with the conclusion of Odenthal et al. (1994) that selective attention skills are related to safe and successful negotiation of demanding traffic situations.

Over the last several years Fisk, Rogers, and their colleagues (e.g., Fisk & Rogers, 2000; Jamison & Rogers, 2000; Mead & Fisk, 1998; Rogers, Campbell, & Pak, 2001; Rogers, Fisk, Mead, Walker, & Cafarella, 1998) and others (e.g., Chamness, Schuman, & Bortz, 1992; Czaja, 1996) have systematically investigated the acquisition and retention of new technology skills by older adults. This work has been motivated largely by an attempt to understand how training methodology should be altered in light of the known changes in cognition that accompany aging.

The current research speaks to several points raised in this literature. First, as Rogers et al. (1996) have pointed out, for older adults interactive training is more effective than other forms of training. In the current research, although both the speed-of-processing and simulator training involved interactive training, the interactive nature of the training was far more extensive in the speed of processing than in the simulator training. This may explain, in part, why the effects of simulator training had largely dissipated by 18 months. Second, as Fisk and Rogers (2000) have noted, older adults, despite declines in cognitive ability that accompany aging, show little if any decline in the performance of many well-learned skills. This is consistent with the findings of the current research, in that many of the driving skills of the older adults were performed at near-perfect levels. However, as Walker, Jain, Fisk, and McGuire (1997) have noted, older adults are slower to make driving decisions (e.g., route selection) than are younger adults, but if they are given sufficient time the quality of their decisions does not decline. It is precisely these decision points, which occur unexpectedly during the drive and require a rapid response, that have been associated with increased crash risk in older adults (e.g., Campbell, 1966; Kline, 1986) and for which the impact of speed-of-processing training can be seen.

In summary, these data add further weight to the accumulating research linking processing speed deficits to driving performance failures by older adults. Furthermore, the current data suggest that these processing deficits can be ameliorated through speed-of-processing training and that this improvement results in safer driving behaviors that are durable for at least 18 months.

ACKNOWLEDGMENTS

A portion of the results of this study were reported at the 1997 Association for Research in Vision and Ophthalmology Conference and abstracted in Roetke, Cisell, and Ball (1997). Daniel Roenker and Karlene Ball are consultants to, and stock holders in, Visual Awareness, Inc., the manufacturer of the speed-of-processing training equipment used in this study. The study was supported by the Edward R. Roybal Center for Research on Applied Gerontology at the University of Alabama at Birmingham and a grant from the National Institute on Aging (R44AG09727).

REFERENCES

Daniel L. Roenker is a professor of psychology at Western Kentucky University and director of the Western Kentucky University Center on Aging. He received his Ph.D. in experimental psychology from Kansas State University in 1973. Gail M. Cassell is the assistant director of the Center on Aging at Western Kentucky University, where she received her M.A. in experimental psychology in 1992. Karlene K. Ball is a professor of psychology and director of the Edward R. Roybal Center for Re-

search on Applied Gerontology at the University of Alabama at Birmingham. She received her Ph.D. in experimental psychology from Northwestern University in 1979. Virginia G. Wadley is an assistant professor of psychology and assistant director of the Edward R. Roybal Center for Research on Applied Gerontology at the University of Alabama at Birmingham, where she received her Ph.D. in medical/clinical psychology in 1997.

Jerri D. Edwards is a postdoctoral fellow at the Edward R. Roybal Center for Research in Applied Gerontology at the University of Alabama at Birmingham, where she received her Ph.D. in developmental psychology in 2000.

Date received: March 14, 2001
Date accepted: January 13, 2003