Published in final edited form as:

Transp Res Part F Traffic Psychol Behav. 2013 September 1; 20: . doi:10.1016/j.trf.2013.07.004.

A simulator evaluation of the effects of attention maintenance training on glance distributions of younger novice drivers inside and outside the vehicle

Gautam Divekar^{a,*}, Anuj K. Pradhan^{b,2}, Kathleen M. Masserang^c, Ian Reagan^d, Alexander Pollatsek^c, and Donald L. Fisher^e

^a Department of Mechanical and Industrial Engineering, University of Massachusetts, Elab 302, Amherst, MA 01003, USA

^b Department of Mechanical and Industrial Engineering, University of Massachusetts, Elab 110, Amherst, MA 01003, USA

^c Department of Psychology, University of Massachusetts, Tobin Hall, Amherst, MA 01003, USA

^d National Highway Traffic Safety Administration, 1200 New Jersey Ave., SE, Washington, DC 20590, USA

^e Department of Mechanical and Industrial Engineering, University of Massachusetts, Elab 220A, Amherst, MA 01003, USA

Abstract

Driver distraction inside and outside the vehicle is increasingly a problem, especially for younger drivers. In many cases the distraction is associated with long glances away from the forward roadway. Such glances have been shown to be highly predictive of crashes. Ideally, one would like to develop and evaluate a training program which reduced these long glances. Thus, an experiment was conducted in a driving simulator to test the efficacy of a training program, FOCAL, that was developed to teach novice drivers to limit the duration of glances that are inside the vehicle while performing an in-vehicle task, such as looking for a CD or finding the 4-way flashers. The test in the simulator showed that the FOCAL trained group performed significantly better than the placebo trained group on several measures, notably on the percentage of within-vehicle glances that were greater than 2, 2.5, and 3 s. However, the training did not generalize to glances away from the roadway (e.g., when drivers were asked to attend to a sign adjacent to the roadway, both trained and untrained novice drivers were equally likely to make especially long glances at the sign).

Keywords

Driver distraction; Eye movements; Nomadic tasks; Younger drivers; Driver training

© 2013 Elsevier Ltd. All rights reserved.

* Corresponding author. Address: Battelle, Center for Human Performance and Safety, 1100 Dexter Ave. N, Suite 400, Seattle 98109, USA. Tel.: +1 413 687 3855; fax: +1 206 528 3555. gautamdivekar2000@gmail.com (G. Divekar). anuj.pradhan@nih.gov (A.K. Pradhan), kmassera@psych.umass.edu (K.M. Masserang), ian.reagan@dot.gov (I. Reagan), pollatsek@psych.umass.edu (A. Pollatsek), fisher@realmail.ecs.umass.edu (D.L. Fisher).

¹ Present address: Battelle, Center for Human Performance and Safety, Seattle, WA 98109, USA.

² Present address: National Institutes of Health, Bethesda, MD 20892, USA.

1. Introduction

A naturalistic study done by the Virginia Tech Transportation Institute in 2006 estimated that 78% of all crashes and 65% of all near crashes observed in the study involved an inattentive or distracted driver (Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006). When estimates are made of particular distracting activities, like cell phone use, the results from epidemiological studies (McEvoy et al., 2005; Redelmeier & Tibshirani, 1997) are at odds with the results of naturalistic studies (Klauer et al., 2006; Olson, Hanowski, Hickman, & Bocanegra, 2009), the former reporting a fourfold increase in crash risk while drivers are on the cell phone, the latter reporting an increase, but not a statistically significant one. The debate about why these differences in the effect of cell phone use on crash risk exist in these two different types of studies is a continuing one (Kidd & McCartt, 2012; Young, 2011). But no one is questioning why texting or other activities which require the driver to take his or her eyes off the road in order to complete the task increases crash risk significantly (Drews, Yazdani, Godfrey, Cooper, & Strayer, 2009; Olson et al., 2009), though there may be some debate about the exact size of the effect.

This problem has not gone unnoticed, beginning with research some 15 years ago (Summala, Lamble, & Laakso, 1998; Summala, Nieminen, & Punto, 1996; Wikman, Nieminen, & Summala, 1998). It has received continuing attention from the research community, including both edited volumes on distraction and driving (e.g., Regan, Lee, & Young, 2009) as well as individual studies of the effect of cognitive distraction (e.g., Muttart, Fisher, Knodler, & Pollatsek, 2007; Strayer, Drews, & Crouch, 2003), in-vehicle distractions (e.g., Chan, Pradhan, Pollatsek, Knodler, & Fisher, 2010; Horrey & Wickens, 2007), and external-to-the vehicle distractions (e.g., Beijer, Smiley, & Eizenman, 2004; Wallace, 2003; Divekar, Pradhan, Pollatsek, & Fisher, 2012) on driver behaviors. The problem is a particularly acute one for younger novice drivers (Braitman, Kirley, McCartt, & Chaudry, 2008; Chan et al., 2010; Klauer et al., 2006; Lee, Olsen, & Simons-Morton, 2006; Lerner & Boyd 2004; McKnight & McKnight, 2003; Wang, Knipling, & Goodman, 1996). For example, in the naturalistic study done by the Virginia Tech Transportation Institute the inattention related crash and near-crash events for younger drivers (age 18–20 years) were estimated to be four times more than more experienced drivers (age 35 and above).

There are many factors that can explain this high inattention or distraction related crash rate for younger drivers. But, the one behavioral attribute that stands out and is the most frequently proposed explanation for the high crash rate for younger drivers is their willingness to look away from the forward roadway for longer durations while performing tasks inside or outside the vehicle as compared to more experienced drivers. A study done over 15 years ago in the field (Wikman et al., 1998) and a more recent study on a driving simulator (Chan et al., 2010) indicate that this type of distraction is a special problem for younger drivers (consistent with the crash results). In the field study conducted by Wikman et al. (1998), the authors found that 46% of inexperienced drivers (mean age 19 years) took glances over 2.5 s inside the vehicle as compared to just 13% of the experienced drivers (mean age 36 years). The simulator study by Chan et al. (2010) found, similarly, that inexperienced drivers performed less safely than did experienced drivers. Specifically, inexperienced drivers performing an in-vehicle task had glances longer than 2.5 s in 45% of the tasks (average age 16.8 years) compared to more experienced drivers who had glances longer than 2.5 s in 10% of the tasks (average age 23.9 years). Both studies also found that there was no significant difference between the groups in terms of total time spent on the task. This means that experienced drivers took smaller but more frequent glances inside the car than younger inexperienced drivers.

Given that the long glances inside the vehicle and away from the forward roadway are a real problem, the question then becomes one of exactly how long it is safe to look away. Estimates of how long is 'too long' have come from two recent studies. First, a simulator study reported by Horrey and Wickens (2007) found that glances of 1.6 s or longer inside the vehicles, while constituting only a relatively small fraction of the total glances (22%), were responsible for the great majority of crashes (86%). Second, consistent with Horrey and Wickens, in the aforementioned naturalistic study reported by Klauer et al. (2006), it was glances longer than 2 s inside the vehicle that increased the crash risk significantly – by about a factor of three.

The willingness of younger inexperienced drivers to look away from the forward roadway might be due to their lack of situational awareness as compared to more experienced drivers (Lee et al., 2006), their failure to recognize the existence of potential hazards due to impoverished mental models (Pradhan, Pollatsek, Knodler, & Fisher, 2009; Underwood, Chapman, Bowden, & Crundall, 2002), or their increased willingness to take risks (Arnett, 2002). These are questions that still need to be addressed. However, in the meantime, there are a number of proposed ways to tackle the problems that distractions create for novice drivers. These include, for example, an outright ban on cell phone use while driving or the setting of standards for design of in-vehicle devices to minimize the time spent on tasks. However, one of the more obvious alternatives to consider is training, especially if the problem is most acutely centered among novice drivers.

Recently a PC-based attention maintenance training program called FOCAL (FOcused Concentration and Attention Learning) was developed and evaluated on a PC (Pradhan et al., 2010). The aim of this program is to produce novice drivers with an operational understanding of how dangerous long glances away from the forward roadway are, and specifically, to train younger novice drivers to limit their glances inside the vehicle to less than 2 s. Towards this end, a recent evaluation of FOCAL training was undertaken in the field (Pradhan et al., 2011). In that evaluation, both novice drivers given FOCAL training and those given Placebo training were asked to navigate streets in the neighborhood of the local high school. They drove in one of two identically equipped vehicles that were controlled by a driving instructor sitting in the front passenger seat. Participants' eyes were tracked throughout the drive. The experimenter explained before beginning the drive that the participant would be asked to perform one of nine secondary tasks. These tasks involved any of the vehicle system controls, the radio, or the materials (CD case, map, or coins) on the center console. The experiment demonstrated that the FOCAL training program significantly reduced the proportion of scenarios in which novice drivers make long glances away from the roadway when drivers are engaged in various in-vehicle tasks. For example, the Placebo trained drivers glanced longer than 2.5 s inside the vehicle in 60.2% of the scenarios whereas the FOCAL trained drivers glanced over this threshold in only 41.9% of the scenarios, a difference of 18.3 percentage points.

Unfortunately, safety considerations make it impossible to evaluate the training program in the field in ways which one can do on a driving simulator. In addition, the data in the field test was somewhat compromised. That is, because of safety considerations, the experimenter sometimes needed to terminate how long the driver in the field performed the in-vehicle task. Thus, it was not possible to get a true measure of the percentage of tasks which would have been performed unsafely without this intervention.

The most promising alternative way to pursue this issue was to run a study on a driving simulator. There are several advantages to doing such. First, we hoped to determine whether FOCAL reduces especially long glances outside the vehicle (e.g., at billboards and electronic message boards) as well as reducing the length of such glances inside the vehicle.

As noted above, there is evidence that long glances at outside distractions are also problematic (e.g., Milloy & Caird, 2011). However, we could not test this safely in the field. Second, we wanted to know whether the training program would work as well when an instructor was not sitting in the front seat as it would when an instructor was sitting in the front seat. It is well known that the crash rate of novice drivers with an adult passenger in the front seat is no different from that of the adult driver (Williams, Preusser, Ferguson, & Ulmer, 1997), but this changes immediately when the novice driver gets his or her solo license (Mayhew, Simpson, & Pak, 2003). Perhaps the training program worked in the field because a passenger was in the front seat or because the FOCAL-trained driver may have felt a need to please the driving instructor and thus took only short glances away from the forward roadway during the secondary in-vehicle tasks (Ambady & Rosenthal, 1996). Regardless, we could not safely perform the experiment in the field without a passenger in the front seat. Finally, we wanted to control precisely the duration of the various in-vehicle tasks in order to understand how drivers' performance varied from task to task. Again, this is not something which we could do safely in the field.

With this in mind, a study of the efficacy of the FOCAL training program was undertaken on a driving simulator where the above safety considerations do not play a role. The distribution of the glance durations of novice drivers trained with FOCAL was compared with the distribution of the glance durations of novice drivers given Placebo training. We hypothesized that for the in-vehicle tasks, FOCAL trained drivers would have fewer glances greater than a threshold of 2 s, would glance longer than 2 s in fewer scenarios, and would have shorter average maximum glances inside the vehicle. We did not have any *a priori* hypotheses about the effect FOCAL training on the glance distributions of drivers performing tasks outside the vehicle.

2. Method

2.1. Participants

The participants were recruited from the Amherst, Massachusetts area. All the participants who participated in the experiment were between the ages of 16–18 years and had held their Junior Operator's license for less than 6 months. Such licenses are issued to Massachusetts' drivers between the ages of 16½ and 18 years old who have had a learner's permit for at least 6 months. The Junior Operator's License allows the driver to operate the vehicle without another adult in the car, though nighttime and passenger restrictions apply. There were 40 participants in the experiment, 23 males and 17 females. Twenty participants each were randomly assigned to either FOCAL or Placebo group. The mean age of participants in FOCAL group was 16.45 and the mean age of participants in Placebo group was 16.40 with standard deviations respectively of 0.68 and 0.59 respectively. There were a total of 11 males and 9 females in the FOCAL group. The mean age of males in the FOCAL group was 16.72 years with a standard deviation of 0.78 and the mean ages of females in this group was 16.11 years with a standard deviation of 0.33. The Placebo group had 12 males and 8 females with the mean ages of males being 16.5 years and mean ages of females was 16.25 years with standard deviations of 0.67 years and 0.46 years respectively.

2.2. Experimental design

The experiment took about 2 h to complete for each participant. In the first hour of the experiment, participants were given a training session. For all participants, the training session began with a PC-based Attention Maintenance Assessment Program (AMAP) which served as pre-test and ended with AMAP as a post-test. Between the pre-test and post-test, the participants were either given FOCAL training or a Placebo training. (Both training programs and AMAP are explained in detail below.) In the second hour of the experiment,

the participants were first familiarized with the simulator by means of a practice drive. Then, in the evaluation drive that followed, they were asked to navigate through a virtual world while performing distraction tasks inside the vehicle (e.g., turning on the high beams) and outside the vehicle (e.g., scanning for a target letter in a sign) (see Section 2.4 for a complete description). The participants drove a total of three drives. In the first drive, participants navigated a two lane visual database (one lane in each direction) populated with vegetation and housing with randomly parked vehicles while performing in-vehicle tasks (see Fig. A. 1a). The posted speed limit was 30 mph. The driver was asked to maintain the speed limit at all times. In the next two drives the participants navigated through the virtual world with the help of a lead vehicle (Fig. A.1b). The participants were instructed to follow the lead vehicle while maintaining a safe following distance between the two. The posted speed limit was again 30 mph. The lead vehicle's speed was matched to the driver's vehicle's speed. The participants were asked to perform out-of-vehicle tasks during these last two drives. These two drives were developed as four lane visual databases (two lanes in each direction) and were populated with vegetation, housing and randomly appearing stationary pedestrians and parked cars.

2.2.1. Attention Maintenance Assessment Program (AMAP)—AMAP is a PC-based interactive attention maintenance assessment tool which contains four filmed videos. Each video is a segment of a real drive that lasts approximately 1 min, is taken from the driver's perspective, and contains driving related elements such as pedestrians and road signs. These videos were used as surrogates for the primary driving task. To simulate the situation of the driver paying attention to the road or to an in-vehicle task, the screen displayed either: (a) the video on the top half of the screen and a black rectangle on the bottom half of the screen (Fig. A.2, top panel) or (b) a map on the bottom half of the screen and a black rectangle on the top half of the screen (Fig. A.2, bottom panel). The participant could toggle between the two views using the "Space" key to switch from the video to the map and the "Enter" key to switch from the map to the video. To simulate the driving task and to assess whether participants were attending to the road, the participants were required to hit the "Enter" key to indicate important driving related events, such as when a road sign, pedestrian or opposing vehicle would pass across the superimposed vertical bars in the video view. (The same key, "Enter", was used for both toggling from the map to the video and indicating important driving related events based on pilot studies which suggested that this two key function was easier for participants to remember than a separate key for all three actions.) Each of the four videos was paired with a unique map. To simulate an in-vehicle task, the participants were required to look for three street names on the map. The street names were displayed before each video began and then again below each map screen. At the end of each video the participants had to indicate the streets that they had found on the map. Not all of the streets that participant were asked to look for on a map were present on that map.

As indicated above, participants in both groups were given AMAP as a pre-test and post-test. Performance was measured by: (a) the times that participants spent on either the video or the map between the key presses and (b) the key presses to indicate the driving related events. The pre-test results of the AMAP tool were used to measure the baseline glance behavior for both groups. The pre-test results of the AMAP tool were also used as input to provide feedback to the FOCAL trained participants in the first phase of their training. AMAP took about 4–5 min to complete.

2.2.2. FOCAL training—The FOCAL training tool had an interactive environment that contained a video of a driving scene and a road map similar to those in AMAP. The presentation format and responses (and coding of responses) were the same as in AMAP.

FOCAL differed in that the map task was somewhat different and that there were the additional training manipulations described below.

FOCAL training consisted of two main sections: (a) feedback and (b) training. In the feedback section, the participants were provided with feedback based on their performance in the pre-test AMAP. (They did not have to make key presses during the feedback section.) The feedback was provided on the AMAP video-map combination in which the participant had the longest single glance at the map (which should have been most illustrative of the dangers of looking away from the forward roadway).

The feedback was given in two phases. In the first phase, the video was played back to the participant with the screen completely blacked out for the exact moments when the participant had originally switched to the map task in the pre-test. For example, if the participant switched to the map for 3.8 s in the pre-test video the screen was blacked out for the same 3.8 s on the video playback. At the end of the playback the participants were instructed via text that they could not see any information on the road when they are looking inside the vehicle. This was done in order to create an appreciation on the part of the participant that they are missing valuable information on the forward roadway when they are looking inside the vehicle. In the second phase of feedback, the same video was played back, but this time with a highly visible timer displayed during the blackout periods. This was done in order to give the participant a feel of how long their glances away from the forward roadway were. This was followed by text that instructed them that they needed to learn not to look away from the forward roadway for more than 2 s.

The training section followed the feedback section. In the training section, the participants had to perform a series of video and map tasks. The map tasks in this case required the participant to identify a cross street which connected the street on which the drivers were traveling and a street that ran roughly parallel to the drivers' direction of travel. The drivers were informed before they began training that there were penalties for: (a) a glance at the map that exceeded the prespecified duration threshold, (b) poor performance on the map search task, and (c) missing essential safety related events in the video. The overall aim of this section was to train the participants to limit their glances to less than 2 s at the map.

The training session had two similar phases, each phase having the same two subparts (in both phases, the first subpart had four video-map tasks and the second subpart had three video-map tasks). In the first phase of training, the time duration threshold mentioned above was set to 3 s (i.e., the participant could spend 3 s looking at the map without receiving a penalty). In the first subpart, four video-map tasks were displayed. The map task stayed on the screen for 3 s and then disappeared. The participant had to repeat the map task only if the connecting street was not identified. In the second subpart, three video-map tasks were displayed. The participant controlled the length of the time the map task was displayed. If this time was greater than 3 s, the participant had to repeat the task. As above, the participant also had to repeat the task if he or she failed to find a street name. Each task in the subpart was repeated for a maximum of three times. This was done in order to limit the total duration of the training program. In the second phase, the time duration threshold was reduced to 2 s. Again, there were two subparts. (Actually, the penalty for poor performance was only based on criteria (a) and (b) above, not on criterion (c), but participants were unaware of this.)

2.2.3. Placebo training—The Placebo training was compiled from the “Massachusetts Registry of Motor Vehicle” driver’s manual (Massachusetts Department of Transportation, 2010). The training contained information regarding various road signs and signals and their implications for the driver. The participants were given 20 min to read through the manual

before they were given a quiz containing ten questions. The participants had to reread and answer a new set of questions if they got one or more questions incorrect on the first quiz. The Placebo training took about the same amount of time as the FOCAL training.

2.3. Simulator evaluation

Once having completed the FOCAL or Placebo training, the participants' glances inside and outside the vehicle both during performance of secondary tasks and during baseline driving were evaluated on a driving simulator using an eye tracker to determine the glance durations and locations.

2.3.1. Driving simulator—The vehicle cab in the Arbella Insurance Human Performance Laboratory mid range driving simulator is a full sized Saturn sedan (see Fig. A.3). A driver operates the controls of the sedan just as he or she would on the road. The visual world is displayed on three screens, one in front of the car and two on each side. Each screen subtends 60° in the horizontal direction and 30° in the vertical direction. As the driver turns the wheel, brakes or accelerates, the roadway that is visible to the driver changes appropriately. The images themselves are updated 60 times a second using a network of four advanced RealTime Technologies Inc. (RTI), simulator servers which parallel process the images projected to each of the three screens using high end multimedia video processors. The image resolution on each screen can be as high as 1024 × 768. The sound system for the simulator consists of three Logitech Dolby 2.1 Surround Sound speakers, two located on the left and right sides of the car and one, a sub-woofer, located in front of the car. The system provides realistic road, vehicle, wind and other noises with appropriate direction, intensity and Doppler shift.

2.3.2. Eye tracker—A portable lightweight eye-tracker (Mobile Eye developed by Applied Science Laboratories) was used to collect the eye-movement data for each driver during the virtual drivers. It has a lightweight optical system consisting of an eye camera and a color scene camera mounted on a pair of safety goggles. The images from these two cameras are interleaved and recorded on a remote system, thus ensuring no loss of resolution. The interleaved video can then be transferred to a PC, where the images are separated and processed. The eye movement data are converted to a crosshair, representing the driver's point of gaze, which is superimposed upon the scene video recorded during the drive. This provides a record of the driver's point of gaze on the driving scene while maneuvering through the virtual world. The remote recording system is battery powered and is capable of recording up to 90 min of eye and scene information in a single session. The eye position is sampled at 25 Hz. Head movement is virtually unlimited with a visual range of 50° horizontal and 40° vertical. The system has an accuracy of 0.58° of visual angle (greater than 95% of the gaze points fall within 0.58° of the target points) and an average vertical precision of 0.09° of visual angle and an average horizontal precision of 0.07° of visual angle (Applied Science Laboratories, 2012).

2.4. Secondary tasks: In-vehicle and out-of-vehicle

2.4.1. In-vehicle secondary tasks—As mentioned earlier, the participants were asked to perform in-vehicle distraction tasks in the first evaluation drive. There were a total of nine in-vehicle tasks: three were related to control of the vehicle and the remaining six were non-vehicle control tasks. The three tasks related to control of the vehicle were, turning the high beams on, activating the emergency flashers, and setting the air conditioner to defrost the front window. The six non-vehicle control related tasks included tuning the radio to a certain channel and frequency, looking for a target street on a map when it was present and when it was not present, looking for a CD in a CD case when it was present and when it was not, and finding 40 cents from the door pocket and depositing it in a box on the passenger seat.

All the in-vehicle secondary tasks were called out using prerecorded audio files activated using virtual coordinate sensors in the simulator drive. This enabled the task to be initiated at exactly the same location for the exact period of time for all participants. The participants were allotted exactly 15 s to perform each task. All the tasks were initiated on straight sections of the drive.

2.4.2. Out-of-vehicle secondary tasks—To evaluate whether FOCAL training had any effect on the glance distribution pattern of drivers with respect to out-of-vehicle distractions, the drivers had to navigate two additional drives. Each of the drives had 12 out-of-vehicle distraction tasks. A lead vehicle braked in one of the distraction tasks in each drive, but the software controlling this vehicle turned out not to be reliable. Thus, we report the data on only 11 of the out-of-vehicle distraction tasks. For the distraction tasks, in one of the drives, the participants were instructed to search for and indicate a target letter that is present on a 5 × 5 letter grid. In this drive, the target letters were “P”, “E” or “X”. In the next drive the participants were asked to look for and indicate the number of times the target was present on the grid. In this case the target letters were “P”, “E”, or “W”. The order in which the participants drove the two out-of-vehicle distraction task drives was counterbalanced across participants.

The letter grid was superimposed on a simulated 10 foot wide by 10 foot high display that was positioned 8 feet from the left or right hand edge of the street (Fig. A.1b, right hand side of the road). Each display would become visible 196.85 feet before the driver encountered the sign. At this point, the display subtended approximately 1.6° of visual angle and its center was 5.1° horizontally from directly in front of the vehicle. If drivers traveled at the posted speed limit (30 mph) the letters on the grid would be visible for 5 s. The grids were populated with letters that had the same visual shape as the target letters (P, E, X or W) to reduce the salience of the target letter if it was present in the grid. In one of the drives, a target letter was present on 3 grids and no target letter was present in the remaining 8 grids. In the other drive, the target letter was present on 7 grids and no letter was present on the remaining 4 grids. The present and absent trials were counterbalanced throughout the experiment.

2.5. Procedure

The study was conducted in the Arbella Insurance Human Performance Lab at the University of Massachusetts Amherst. After the initial briefing about the study and its duration, the participants filled out a demographic questionnaire and then were given their respective training, either FOCAL or Placebo. Once the training was done the participants had to fill out a simulator sickness questionnaire (SSQ, Kennedy, Lane, Berbaum, & Lilienthal, 1993). The aim of the questionnaire was to screen the participants for any indication of motion sickness based on the previous exposure to either virtual environments or equipment motion that has been associated with nausea such as roller coasters, boats, planes and trains. The participants were given brief instructions about driving the simulator and then they drove a stand-alone practice drive to familiarize themselves with the handling of the simulator cab. Once the practice drive was completed, the participants were fit with the mobile eye tracker and then the experimenter performed the eye tracker calibration. The eye tracker set up and calibration took about 5–7 min. After the calibration was done the experimenter explained that the secondary tasks would involve both vehicle control systems and task specific materials like a CD case, a map and some coins that were placed in the car. The participants were told that the secondary task would be called automatically through the audio system in the simulator. The posted speed in all three drives was 30 mph. Participants were told to maintain the posted speed. In addition in the second and third drive, the participants were instructed to follow a lead vehicle while maintaining a safe distance. The

speed of the lead vehicle was tied to the participant's vehicle's speed; this arrangement allows the participant to drive at his or her natural speed with respect to the posted speed limit and also helps to reduce any influence that the lead vehicle might have on participant's driving speed. Once the participant indicated that all instructions given to him or her were clear the evaluation drive was initiated.

3. Results

The primary unit in the analysis was the *task*. Eye movement data were only scored during the tasks. Moreover, the primary unit of eye movement behavior was the *glance*. This was defined somewhat differently than in other fields using eye movement measures, such as reading, where a glance is defined as the sum of fixation durations on a given object, such as a word (e.g., Rayner, 1998). Instead, a glance is defined here as the total amount of time between when the eyes leave the forward roadway and when they return to the forward roadway, but does not include the leading or trailing saccades (sometimes leading saccades are also included in the glance time, e.g., see ISO, 2002a, 2002b; Society of Automotive Engineers, 2001). For the within-vehicle tasks that will be discussed first and to which the most attention will be paid, this time undoubtedly includes saccadic eye movements inside the vehicle and may often include fixations on more than one object or area within the vehicle. However, it seems like the most ecological index of the length of time that vision (and attention) is not directed towards the forward roadway. In the out-of-vehicle tasks that are discussed later, the glance includes the period of time in which the eyes are directed at the object to the side of the forward roadway (but again not the leading or trailing saccades). These glances are thus, in one sense, more like the glances defined in reading and static scene viewing in that they are recording the time that the eyes are pointed at a given object or area. However, they also differ, because the eyes are actually involved in smooth pursuit movements (possibly with some small saccades) as the object is moving with respect to the driver. Below, the results of this study are discussed along with the relation between the results in this study and a related field study (Pradhan et al., 2011) as well as the relation that obtains on the one hand between trained and untrained novice drivers in this study and on the other hand between experienced and untrained novice drivers in a related simulator study (Chan et al., 2010).

3.1. Analysis of in-vehicle glances during in-vehicle tasks

3.1.1. Proportion of tasks in which glances greater than threshold—As indicated above, there were nine in-vehicle tasks. Two primary indices of glance behavior were computed. Both were indices of the frequency of long glances, but they assessed different frequencies. The first measure is the proportion of tasks in which there was a glance greater than a threshold of x seconds. As can be seen from Fig. A.4, there were large differences between the two groups for almost all threshold values x between 1.5 s and 6 s. However, the statistical analyses will concentrate on threshold values of x equal to 2, 2.5, and 3 s, as these seem most closely related to crash rates (Horrey & Wickens, 2007; Klauer et al., 2006). As can be seen in Table A.1, there were large differences between the two groups for each of the above three threshold values, $t(38) = 3.297, 4.130, 2.598, p < .005, .001, .02$, respectively. The differences between the proportions ranged from about .25 to .35. However, if one considers the comparison in terms of ratios, FOCAL training reduced the number of tasks in which there was a glance over 2 s by about 35% and the number of tasks in which there was a glance over 2.5 s or 3 s by over 50%. Note that although the above statistical tests are clearly not independent, the decision was made to report tests for each value because these are the most critical thresholds. The point here is that regardless of which particular threshold one decides is most important, the training had an effect.

3.1.2. Proportion of glances over threshold—The second primary measure, the proportion of glances that were over a given threshold, showed similarly large effects (see Fig. A.5 and Table A.2). The absolute values of the numbers are smaller, which is a logical consequence of what is being measured, and thus the absolute values of the differences were smaller. Nonetheless, the differences were highly significant for the 2 s, 2.5 s, and 3 s values, $t(38) = 3.735, 3.129, 3.186, p < .001, .001, .005$. In terms of ratios, the FOCAL training reduced the proportion of long glances about 50% for the 2 s and 2.5s thresholds and over 60% for the 3 s thresholds.³ Furthermore, because of some concern of the slight gender imbalance in our sample, we reran the analysis for the crucial 2 s threshold using the R-analysis program making gender a specific variable. The effect of FOCAL training was still highly significant, $F(1,36) = 15.48, p < .001$. There was a gender main effect, $F(1,36) = 6.89, p < .05$, with women having fewer long glances than men; however, there was no group by gender interaction ($F < 1$), so that this slight gender imbalance is of little concern in any of the above analyses.

3.1.3. Average maximum glance durations and average glance durations by task—Also, consistent with our prior studies (Chan et al., 2010; Pradhan et al., 2011), another index computed was the average over the nine tasks of the maximum glance duration in each task. The average maximum times for the Placebo and FOCAL group were 3.014 s and 2.404 s, respectively, and the difference was significant, $t(38) = 2.712, p < .01$.

Finally, we also looked at the average glance duration for each of the nine tasks to determine whether FOCAL was having an effect in some tasks, but not in others. In fact, the average glance duration in all tasks was shorter with the FOCAL trained participants than it was with the Placebo trained participants (average glance duration across all task for FOCAL trained participants was 1.31 s as compared to 1.63 s for Placebo trained participants).

3.1.4. Near and far transfer—One concern we had in reporting the overall data was whether the training that used map tasks transferred well to the other tasks used in the test in the driving simulator. As a result, the data for the two map tasks and for the seven other tasks used in the simulator test were analyzed separately. The basic finding was that the advantage of the FOCAL group over the Placebo group was about the same for the map tasks (i.e., tests of near transfer) and for the other seven tasks (i.e., tests of far transfer) for both of the above measures. First, for the proportion of tasks in which the maximum glance was over x seconds, if anything, the differences were larger by this measure for the non-map tasks (see Table A.3). For the non-map tasks, the differences between the placebo and FOCAL groups were significant for the 2 s, 2.5 s, and 3 s thresholds, $t(38) = 3.344, 4.237, 2.815, p < .005, .001, .01$, respectively. The differences were numerically smaller for the map tasks and only significant at the 2.5 s threshold, $t(38) = 1.674, 2.299, 1.648, p < .20, .05, .20$, respectively for the 2 s, 2.5 s, and 3 s thresholds. (However, it should be remembered that the data for the map tasks were more variable as there were only two observations for each participant.)

For the proportion of glances measure, there were significant differences between the two groups at the 2 s, 2.5 s, and 3 s thresholds for the both non-map tasks (see Table A.4), $t(38) = 3.392, 2.951, 3.381$, respectively, $ps < .005$, and for the map tasks, $t(38) = 3.428, 2.669, 2.039, p < 0.001, 0.025, 0.05$, respectively. By the former measure (proportion of scenarios), the sizes of the differences between groups were larger for the non-map tasks, but by the latter measure (proportion of glances), they were somewhat larger for the map tasks. Given

³We also computed percent of gazes over a threshold by simply pooling all the within-vehicle glances for each participant and computing the proportion of them over 2 s, 2.5 s, and 3 s. The results were quite similar to the above.

the difference in pattern between the two measures and the fact that the map task data is quite variable, one cannot make a definitive statement about whether the far transfer effect is larger or smaller than the near transfer effect. However the data clearly indicate that, by both measures, there was robust transfer of the FOCAL training (which only involved map tasks) to the other seven tasks used in the simulator test.

3.1.5. Total number of glances per task—It is clear from Figs. A.4 and A.5 that training had an effect on the distribution of glances of drivers inside the vehicle: both the proportion of scenarios and the proportion of glances longer than some threshold value were larger at any given threshold for the Placebo group than it was for the FOCAL group. However, if the FOCAL group was making fewer glances inside the vehicle than the Placebo group, then this could easily inflate the difference between the FOCAL and Placebo groups: the more glances made during any given scenario the larger is the likelihood that one of the glances will be over some threshold. In fact, as can be seen in Table A.5, the FOCAL group actually made more glances inside the car than the Placebo group, although not significantly so. Thus, the differences between the FOCAL and Placebo groups in the proportion of scenarios in which the maximum is above some threshold value cannot simply be explained by differences in the number of glances inside the vehicle.

3.1.6. Performance on the in-vehicle tasks: Speed-accuracy trade-offs—Finally, one can ask how well the two groups performed on the various in-vehicle tasks and whether there is any evidence of a glance performance-accuracy tradeoff. The accuracy data indicated that there was little difference between the groups in accuracy so that the superior performances for the FOCAL group in maintaining short glance durations in the vehicle was not at the expense of doing the in-vehicle tasks. In this analysis, the accuracy of the two search tasks in which the object was not present was not scored, as it is not clear when a participant reports that the object was absent that this is because of an exhaustive search or that they did not find it in an incomplete search. Tasks were scored in a binary fashion, as giving partial credit turned out to be too difficult. Performance in neither group was particularly high as these are difficult tasks; the average percent correct for the Placebo and FOCAL groups was 55.0% and 52.3%, respectively. (“Chance” is difficult to assess, but it is much less than 50%.) This 2.7% difference was far from significant, $t < 1$.

3.2. Analysis of out-of-vehicle glances during out-of-vehicle tasks

As indicated above, the participants did the out-of-vehicle distraction tasks in blocks 2 and 3 of the post-training test session. Data are available only from 16 of the FOCAL-trained group participants and 19 of the placebo-trained participants as the other five participants were excused from further testing as they were experiencing simulator sickness. The analyses below also do not include a few trials (three for the FOCAL group and six for the Placebo group) in which there were no glances at the grid. (The participants gave an incorrect response on all these tasks.)

The basic finding is that there was virtually no training effect on the out-of-vehicle distraction tasks, and the one significant effect was in the “wrong” direction. First, consider as the measure the proportion of tasks in which the maximum glance was greater than a fixed threshold (see Fig. A.6 and Table A.6). There was virtually no difference between the groups at the critical 2 s threshold and only a suggestion of an effect for the 2.5 s and 3 s thresholds, $|t(33)| < 1$, $t(33) = 1.211, 1.431, p = .23, .16$, respectively. However, it should be noted that the difference of .114 between the groups at the 3.5 s threshold was significant, $t(33) = 2.050, p < .05$. The FOCAL group also had a somewhat longer average maximum glance duration than the Placebo group (3.365 s vs. 3.216 s), but the difference was not close to significant, $t(33) = 1.509, p = .14$.

The picture was similar for the proportion of glances measure (see Table A.7) with the Placebo group having slightly fewer long glances than the FOCAL group. As with the above measure, none of the differences were close to significant, $t(35) = 1.229, 1.609, 1.619, p = .23, >.12, >.11$, respectively, for the 2 s, 2.5 s, and 3 s thresholds. The reason for the similarity between the measures is made apparent when one examines the average number of glances per task. The average numbers of glances per task for the FOCAL and Placebo groups were 1.034 and 1.071, and the difference was not significant, $t(35) = 1.790, p = .08$. Moreover, as the average number of glances per task indicates, there was virtually no inclination for either group to apportion their glances. Instead, for most out-of-vehicle tasks, they accomplished the task with one single glance at the out-of-vehicle display and were quite accurate – 92.4% and 91.1% correct for the Placebo and FOCAL groups, respectively, $|t| < 1$. It should be noted that, for both groups, the total time they spent on the tasks was appreciably less than the total amount of time the sign was visible (5 s), as the average total time per task was 3.389 s for the FOCAL group and 3.257 s for the Placebo group. (The average maximum times per task were 3.365 s for the FOCAL group and 3.216 s for the Placebo group.) Thus, they were not forced to look continually at the grid in order to perform the task. The difference between groups was not significant on either measure, $t(33) = 1.390, 1.509, p = .17, p > .14$, respectively.

As is obvious, the data for the out-of-vehicle tasks is quite at variance with that in the in-vehicle tasks. First, for both groups, the proportion of tasks in which the maximum glance duration was considerably greater than the 2, 2.5, and 3 s thresholds was greater for the out-of-vehicle tasks, $ts(15) > 5.8, ps < .001$ for the FOCAL group and $ts(18) > 3.2, ps < .005$ for the Placebo group. Moreover, the pattern of data is obviously different; instead of the FOCAL group having many fewer long glances, there was a slight tendency in the opposite direction for the out-of-vehicle tasks. The group by in-vehicle vs. out-of-vehicle interaction was significant for the 2, 2.5, and 3 s thresholds, $ts(15) > 3.0, ps < .005$. It thus appears that the in-vehicle task training is viewed as not relevant to the out-of-vehicle tasks. Another piece of data that corroborates this conclusion comes from correlating in-vehicle and out-of-vehicle performance within groups. For the Placebo group, there were sizeable and consistent correlations between the average probability of exceeding the 2, 2.5, and 3 s thresholds between the two types of tasks, 0.253, 0.301, and 0.473, respectively, $t(17) = 1.080, 1.299, 2.211, p > .20, p > .20, p < .05$. In contrast, for the FOCAL group, the respective correlations were small and inconsistent in sign: $-0.151, 0.101, -0.026 (|t|s < 1)$. This difference suggests that the variability in the Placebo group was largely accounted for by the individual participant's prior inclination to take long glances, whereas for the FOCAL group, the individual differences in the two types of tasks were likely to be due to different factors. What seems most plausible is that the individual differences in the out-of-vehicle tasks for the FOCAL group were due to the participant's prior inclination to take long glances whereas the individual differences in the in-vehicle tasks were due to the participant's attentiveness to the FOCAL training.

3.3. AMAP performance

The performance of the two groups on the AMAP computerized test that they took both before and after the training session (either FOCAL or Placebo training) was also examined. There are two reasons for presenting these data. The first is to see whether the groups performed comparably on the pretest to double-check whether the groups were indeed equated. To keep things simple, only one type of measure is used here: the percent of glances away from the forward roadway that were above either 2, 2.5, or 3 s (where a "glance" here is the time between a key press to expose the map and the subsequent key press to expose the forward roadway).

Unfortunately, there did appear to be a difference between groups on the pretest, with the FOCAL trained group performing somewhat better. The proportion of glances greater than 2, 2.5, and 3 s were 0.742, 0.624, and 0.494 for the Placebo group and 0.653, 0.550, and 0.441 for the FOCAL group. Although none of these differences were close to significant, $t(38) = 1.273$, $p = .21$, $t < 1$, $t < 1$, this is some cause for concern. As a result, an analysis of covariance was run on the data presented in Tables A.1 and A.2 using pretest scores as the covariate. For the proportion of tasks in which the maximum glance was over 2, 2.5, and 3 s, the sizes of the differences reported in Table A.1 decreased by about one quarter and only the first two were significant, $t(37) = 2.576$, 3.224 , 1.806 , $p < .02$, $p < .005$, $p = .08$. The effect of the covariate was even less for the differences reported in Table A.2, with the sizes of the differences being reduced only by about one-tenth and all three effects still were significant, $t(37) = 3.322$, 2.856 , 2.984 , $ps < .01$.

Also of interest was the size of the post-test pretest difference on the AMAP test for the two groups. The differences for the proportion of glances away from the forward roadway that were above either 2, 2.5, or 3 s were 0.089, 0.074, and 0.053, for the Placebo group, whereas the respective differences for the FOCAL group were 0.837, 0.888, 0.814, $F(1,38) = 133.7$, 137.8 , 100.4 , $ps < .001$. Although it is not surprising that the FOCAL group improved hugely, it was surprising that their post-test performance was almost perfect: the mean proportion of off-road glances greater than 2 s was .033 and the median proportion of off-road glances greater than 2 s was 0. Thus, essentially all but one outlying FOCAL participant learned their lesson perfectly on the PC; however, transfer to the driving simulator was less than perfect.

4. Discussion

As indicated in the introduction, it has been determined that a major cause of crashes on the road is inattention to the forward roadway for periods of 2 s or more and that such inattention also plays a role in the elevated crash rates for novice drivers. As a result, the major objective of this experiment was to determine whether our program, FOCAL, would be effective in training novice drivers not to glance away from the forward roadway for such extended periods of time when no adult passenger was present in the front seat of the vehicle. The evaluation in the present experiment on a driving simulator indicated that the answer was definitely “yes” for the in-vehicle tasks. The FOCAL-trained group made significantly fewer glances of greater than 2, 2.5, or 3 s than the placebo-trained group. This was true even when controls were implemented for possible inadvertent prior differences in such tendencies. The second major finding was that the training, which was on map tasks, extended just as well to all of the other in-car tasks tested on in the simulator, so that the training appeared to be clearly understood as “limit the duration of your in-vehicle glances” for a wide range of such tasks. The third major finding was that this training did not transfer at all to the out-of-vehicle tasks. It thus appears that the trainees did not view these tasks as being exemplars of the same concept as the in-vehicle tasks.

We should discuss briefly why we believe FOCAL works as a training program. It has been our experience with the design of other training programs, both for novice drivers (hazard anticipation; Pollatsek, Narayanaa, Pradhan, & Fisher, 2006) and older drivers (Romoser & Fisher, 2009), that there are three elements critical to the success of such programs. The participant (a driver in this case) who is being trained to perform a given skill needs first to make a mistake which he or she finds believable. The participant then needs to be informed of this mistake and told how to mediate the mistake. And finally the participant needs to master the skill so that the mistake no longer occurs. FOCAL, like our other driver training programs, contains these three critical elements. We are speculating, but did not test, whether it is these three critical elements which make FOCAL effective.

4.1. Comparison of simulator and field studies

Four additional questions are of interest. The first question of interest is how the current results compared to those in the field study of Pradhan et al. (2011). As the field study did not contain out-of-vehicle distraction tasks, this comparison is restricted to the in-vehicle tasks. However, because the out-of-vehicle tasks in the simulator study always came in blocks 2 and 3 of the test in the experiment, we think there is little chance that their presence in the present study would have been a contributor to any differences found between the two experiments in the in-vehicle tasks. The answer to the first question is that the training effects were larger in the present simulator experiment than in the field study. We will focus on one measure, the percent of scenarios in which there was a glance greater than threshold. The differences between the FOCAL and placebo trained groups on the percent of scenarios in which there was a glance greater than threshold in the field study were 11.4%, 15.4%, and 11.5% at the 2 s, 2.5s, and 3s thresholds; the comparable differences in the current study were 22.0%, 31.0% and 20.6% (see Table A.1). (Note that the differences were similar on the second measure, the percent of glances greater than threshold measure.)

One possible reason for the difference is that three vehicle-control tasks in the field study presented little useful data as the participants were familiar with the vehicles and thus completed the task very quickly with little need to glance away from the roadway. In contrast, as the participants in the current study were not familiar with the vehicle, the vehicle control tasks in the present study were more difficult and required allocation of attention. When these three tasks were taken out of the analysis, the differences between the studies was somewhat smaller: the differences between groups in the field study were 16.5%, 18.3%, and 11.5% for the 2 s, 2.5 s, and 3 s thresholds, whereas the comparable values for the simulator study were 21.7%, 29.2%, and 19.2%. A second possible reason for the smaller effects in the field study is that, although the participants were randomly assigned to groups, it appeared that the FOCAL-trained group participants, on average, were somewhat more prone to take longer glances as indicated by their scores on the AMAP pretest. A third possible reason for the smaller effect in the field study is the one mentioned in the introduction: the presence of the driving instructor in the front seat who could help out in case of trouble by stepping on the brake may have made the FOCAL group in the field study somewhat more likely to take chances in spite of the training. However, on the whole, the results in the two studies were reasonably similar.

The above comparisons raise several related questions, specifically the validity of testing training programs such as FOCAL on the simulator, and, more generally, the place of driving simulators in evaluating driving training. Our results in this experiment and the related field study suggest that there is definitely a role for driving simulators. Although the results of the two studies do not match perfectly, they match pretty well given that the studies were run in different states (Massachusetts vs. New Jersey) and likely had somewhat different populations of participants (although in both experiments, participants were randomly assigned to either the FOCAL or Placebo training group). In addition, the virtual environments through which the drivers navigated in the present study were not purposely made to resemble the environment of the field study. Moreover, perhaps of most significance to a comparison of the two studies, the videos used in training were filmed in the Amherst, Massachusetts area and thus may have had a greater impact on the participants in the current study.

This similarity between the effects of training programs in the field and on the driving simulator holds for other training programs as well. For example, beyond the FOCAL evaluation discussed above, we have developed another training program, RAPT, that trains novice drivers to anticipate unseen potential hazards (Pollatsek et al., 2006). (This is another documented major cause of increased accidents for novice drivers, e.g., see McKnight &

McKnight, 2003.) In the case of RAPT training, the results were quite similar in field and simulator evaluations (including the sizes of the effects, Pradhan et al., 2009). Thus, the data we have collected so far indicates that the data from simulator evaluation appears to be quite consonant with those from field evaluation. As a result, we think there is a real place for evaluation of training programs on a simulator – both because one has a greater control over the experimental manipulations and because one is not exposing the participants to any danger.

4.2. Comparison of trained novice with experienced drivers

The second question of interest is how the effectiveness of training in the present simulator study compared to differences between experienced drivers and similar inexperienced drivers that were evaluated in the same scenarios on the same simulator (Chan et al., 2010). The two studies are not exactly comparable, however. First, the novice drivers in the Chan et al. study did not undergo any training session, so in this respect, they were not exactly comparable to the Placebo training group in the present study. Second, and perhaps more importantly, the Chan et al. study only had five in-vehicle tasks rather than the nine tasks in the present study. This might have made the participants more careful about looking away from the road, and indeed the novice drivers performed appreciably better in the Chan et al. study than the Placebo control group in the present study. Nonetheless, it seems of interest to compare the difference between the FOCAL and Placebo groups in the present study with the difference between the more experienced drivers and novice drivers in the Chan et al. study to get a “benchmark” for the size of the training effect in the present study.

For the percentage of tasks measure, the differences between groups for the five in-vehicle tasks used in Chan et al. were quite comparable at the three threshold values: 36.7%, 35.0%, and 26.6% for Chan et al. and 22.1%, 31.0%, and 20.5% in the present study at the 2, 2.5, and 3 s thresholds, respectively. However, the overall percentage of long looks was appreciably lower in the Chan et al. study. Thus if the ratio of the percentages in the two groups was used as the measure of “difference between groups”, the effect sizes in the Chan et al. study would be somewhat bigger.

Even with this uncertainty, however, it appears that FOCAL training turned inexperienced drivers into a pretty good approximation of experienced drivers in avoiding long looks. Another interesting comparison between the two studies is the pattern in the out-of-vehicle distraction tasks. Just as training had no effect on performance in the out-of-vehicle tasks in the current experiment, there was no difference between the experienced and inexperienced drivers in Chan et al. In Chan et al., as in the present study, both groups made many long glances at the out-of-vehicle distractor, and the differences in percentage of tasks in which there was a long glance (novice minus experienced groups) were small and not close to significant: 0.9%, 6.0%, and 1.6% for the 2, 2.5, and 3 s thresholds, respectively. Hence, it appears that drivers do not generalize the lesson that one should avoid long glances in the vehicle away from the forward roadway to avoiding long glances outside the vehicle – whether this lesson is learned in our training session or from years of experience on the road. We know that external distractions increase drivers’ risk of crashing. If it can be shown that long glances are responsible for this increase in risk, this suggests that a new training regimen is needed – for both novice and experienced drivers – to make them aware that long glances away from the forward roadway to external distractions should also be avoided.

4.3. Out-of-vehicle tasks

The chief remaining issue is why there is no generalization from training to the out-of-vehicle tasks. This appears to be due to a failure to connect the two types of activities since, as we argued above, the same kind of disconnect occurred with experienced drivers (Chan et

al., 2010). Our best guess is that drivers fail to understand how dangerous long glances outside the vehicle can be because drivers have the illusion that they can monitor the road sufficiently when looking off to the side and thus that there is no need to come back to monitor it periodically (Divekar et al., 2012). In contrast, in the in-car tasks, it is highly unlikely that someone is under the illusion that they are monitoring the road, and although they may initially be under the illusion that they “can get away with” doing so, both our training program and general experience in driving helps to dispel the latter illusion.

4.4. Limitations

In addition to the above, we should mention several limitations of the study. First, we did not ask about participants about their use of drugs, level of education or previous crashes, nor did we measure their near visual acuity. Although we randomly assigned participants to the Focal and Placebo groups, there still could exist differences between the groups on these measures. Second, the sample was one of convenience and as such may not be representative of the larger population of novice drivers. Third, we simply do not know whether the training program will reduce the glance durations of the novice drivers when they are alone or with other teen passengers while they are driving on the open road. Fourth, we did not use all possible measures of the effect that interaction with in-vehicle devices has on the distribution of glances. For example, measures such as the percentage of fixations on road center and the standard deviation of the duration of the glances made inside the vehicle have been shown to be sensitive to a driver's interaction with in-vehicle tasks (Victor, Harbluk, & Engstrom, 2005). Instead, we focused on measures with known relations to crashes (e.g., the percentage of glances greater than some threshold). Fifth, we do not know whether the reductions in glance durations will actually lead to reductions in crashes. Sixth, we do not know how long the effects of FOCAL training will endure. Future studies should address these limitations.

5. Conclusion

In sum, the present experiment demonstrates that the FOCAL training program significantly reduced the duration of glances away from the road in a simulator environment even in the absence of an instructor in the front seat. This suggests that the effect of FOCAL training observed on the road was not simply the consequence of the trained novice drivers behaving well in the presence of an instructor. Indeed, the effect of FOCAL training was larger in the present study, although the (admittedly weak) tests of this difference across experiments indicated that it was not significant. The differences in the present experiment mirror the difference in glance durations between experienced and novice drivers in a simulator environment and thus indicate that the approximately 50 min training program, though not perfect, is doing quite a reasonable job. The one negative result is that the FOCAL training program, which only trained people on limiting glance durations while performing in-vehicle tasks, did not generalize to out-of-vehicle tasks. However, this failure also mirrors the results from Chan et al. (2010) which indicate that experience has also failed to alert drivers to the hazards of looking away from the forward roadway for extended periods of time when attending to something outside the vehicle. We hope, in the future, to be able to extend our training program to get novice (and possibly also experienced) drivers to understand that such long glances are dangerous and should be avoided.

Acknowledgments

This research was funded by grants from the National Highway Traffic Safety Administration and the National Institutes of Health Grant Number 1R01HD057153 to Donald L. Fisher and Alexander Pollatsek. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the NHTSA or NIH.

References

- Ambady, N.; Rosenthal, R. Experimenter effects.. In: Manstead, ASR.; Hewstone, M., editors. Blackwell encyclopedia of social psychology. Blackwell Publishing; Malden, MA: 1996.
- Applied Science Laboratories. Mobile eye: Statement of precision and accuracy example (technical report). Applied Science Laboratories; Bedford, MA: 2012.
- Arnett JJ. Developmental sources of crash risk in young drivers. *Injury Prevention*. 2002; 8(Suppl. 2):ii17–ii21. (discussion ii21–ii23). [PubMed: 12221026]
- Beijer D, Smiley A, Eizenman M. Observed driver glance behavior at roadside advertising signs. *Transportation Research Record*. 2004; 1899:96–103.
- Braitman KA, Kirley BB, McCartt AT, Chaudry NK. Crashes of novice teenage drivers: Characteristics and contributing factors. *Journal of Safety Research*. 2008; 39:47–54. [PubMed: 18325416]
- Chan E, Pradhan AK, Pollatsek A, Knodler MA, Fisher DL. Are driving simulators effective tools for evaluating novice drivers' hazard anticipation, speed management, and attention maintenance skills? *Transportation Research Part F: Traffic Psychology and Behaviour*. 2010; 13:343–353.
- Divekar G, Pradhan AK, Pollatsek A, Fisher DL. Effect of external distractions. *Transportation Research Record: Journal of the Transportation Research Board*. 2012; 2321:15–22.
- Drews FA, Yazdani H, Godfrey CN, Cooper JM, Strayer DL. Text messaging during simulated driving. *Human Factors*. 2009; 51:762–770. [PubMed: 20196300]
- Horrey, WJ.; Wickens, CD. *Transportation Research Record*. Transportation Research Board of the National Academies; Washington, DC: 2007. 2018. In-vehicle glance duration distributions, tails, and model of crash risk.; p. 22-28.
- ISO. International Standard 15007-1: Road vehicles – Measurement of driver visual behaviour with respect to transport information and control systems – Part 1: Definitions and parameters. 2002a Author.
- ISO. International Standard 15007-1: Road vehicles – Measurement of driver visual behaviour with respect to transport information and control systems – Part 2: Equipment and procedures. 2002b Author.
- Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*. 1993; 3(3):203–220.
- Kidd, DG.; McCartt, AT. Review of “cell phone use and crash risk: Evidence for positive bias”. Young, RA., editor. Insurance Institute for Highway Safety; Arlington, VA: 2012.
- Klauer, SG.; Dingus, TA.; Neale, VL.; Sudweeks, JD.; Ramsey, DJ. The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. National Highway Traffic Safety Administration; Washington, DC: 2006.
- Lee SE, Olsen ECB, Simons-Morton B. Eye glance behavior of novice teen and experienced adult drivers. *Transportation Research Record*. 2006; 1980:57–64.
- Lerner, N.; Boyd, S. Task report: On-road study of willingness to engage in distracting tasks. Report jointly funded under National Highway Safety Administration (Contract DTNH22-99-D-07005, Task Order Number 7) and National Institute of Child Health and Human Development (Contract GS-23F-8144H). Westat; Rockville: MD: 2004.
- Massachusetts Department of Transportation. [31.08.10] Welcome MassDOT – Registry division. 2010. <www.massdot.state.ma.us/rmv>
- McKnight JA, McKnight SA. Young novice drivers: Careless or clueless. *Accident Analysis and Prevention*. 2003; 35:921–925. [PubMed: 12971927]
- Mayhew DR, Simpson HM, Pak A. Changes in collision rates among novice drivers during the first months of driving. *Accident Analysis and Prevention*. 2003; 35:683–691. [PubMed: 12850069]
- McEvoy SP, Stevenson MR, McCartt AT, Woodward M, Haworth C, Palamara P, et al. Role of mobile phones in motor vehicle crashes resulting in hospital attendance: A case-crossover study. *British Medical Journal*. 2005; 331:428–430. [PubMed: 16012176]

- Milloy, SL.; Caird, JK. External driver distractions: The effects of video billboards and wind farms on driving performance.. In: Fisher, D.; Rizzo, M.; Caird, J.; Lee, J., editors. Handbook of driving simulation for engineering, medicine and psychology. CRC Press; Boca Raton, FL: 2011.
- Muttart J, Fisher DL, Knodler M, Pollatsek A. Driving simulator evaluation of driver performance during hands-free cell phone operation in a work zone: Driving without a clue. Transportation Research Record. 2007; 2018:9–14.
- Olson, RL.; Hanowski, RJ.; Hickman, JS.; Bocanegra, J. Driver distraction in commercial vehicle operations. Federal Motor Carrier Safety Association; Washington, DC: 2009. (report no. FMCSA RRR-09-042)
- Pollatsek A, Narayana V, Pradhan A, Fisher DL. The use of eye movements to evaluate the effect of PC-based risk awareness training on an advanced driving simulator. Human Factors. 2006; 48:447–464. [PubMed: 17063961]
- Pradhan, AK.; Masserang, K.; Pollatsek, A.; Divekar, G.; Romoser, M.; Regan, I., et al. Training attention maintenance: A first step.; Transportation Research Board 89th Annual Meeting 2010. 2010 paper 10-0391
- Pradhan AK, Divekar G, Masserang K, Romoser M, Zafian T, Blomberg RD, et al. The effects of focused attention training (FOCAL) on the duration of novice drivers' glances inside and outside the vehicle. Ergonomics. 2011; 51:917–931. [PubMed: 21973003]
- Pradhan AK, Pollatsek A, Knodler M, Fisher DL. Can younger drivers be trained to scan for information that will reduce their risk in roadway traffic scenarios that are hard to identify as hazardous? Ergonomics. 2009; 62:657–673. [PubMed: 19296315]
- Rayner K. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin. 1998; 124:372–422. [PubMed: 9849112]
- Redelmeier DA, Tibshirani RJ. Association between cellular-telephone calls and motor vehicle collisions. New England Journal of Medicine. 1997; 336:453–458. [PubMed: 9017937]
- Regan, MA.; Lee, JD.; Young, K. Driver distraction: Theory, effects, and mitigation. CRC Press, Taylor Francis Group; Boca Raton, FL: 2009.
- Romoser M, Fisher DL. The effect of active versus passive training strategies on improving older drivers' scanning for hazards while negotiating intersections. Human Factors. 2009; 51:652–668. [PubMed: 20196291]
- Society of Automotive Engineers. Measurement of driver visual behavior using video based methods (definitions and measurement) (SAE recommended practice J2396). Society of Automotive Engineers; Warrendale, PA: 2001.
- Strayer DL, Drews FA, Crouch DJ. A comparison of the cell phone driver and the drunk driver. Human Factors. 2003; 48:381–391. [PubMed: 16884056]
- Summala H, Nieminen T, Punto M. Maintaining lane position with peripheral vision during in-vehicle tasks. Human Factors. 1996; 38:442–451.
- Summala H, Lambale D, Laakso M. Driving experience and perception of the lead car's braking when looking at in-car targets. Accident Analysis and Prevention. 1998; 30:401–407. [PubMed: 9666236]
- Underwood G, Chapman P, Bowden K, Crundall D. Visual search while driving: Skill and awareness during inspection of the scene. Transportation Research Part F. 2002; 5:87–97.
- Victor TW, Harbluk JL, Engstrom JA. Sensitivity of eye-movement measures to in-vehicle task difficulty. Transportation Research Part F. 2005; 8:167–190.
- Wallace, B. External-to-vehicle driver distraction. Transport Research Series (Transport Research Planning Group). Scottish Executive Social Research; Edinburgh: 2003.
- Wang, JS.; Knipling, RR.; Goodman, MJ. The role of driver inattention in crashes: New statistics from the 1995 crashworthiness data system.. 40th Annual proceedings of the association for the advancement of automotive medicine; Vancouver, British Columbia. 1996.
- Wikman A, Nieminen T, Summala H. Driving experience and time-sharing during in-vehicle tasks on roads of different width. Ergonomics. 1998; 41:358–372.
- Williams AF, Preusser DF, Ferguson SA, Ulmer RG. Analysis of the fatal crash involvements of 15-year-old drivers. Journal of Safety Research. 1997; 28:49–54.
- Young RA. Cellphone use and crash risk: Evidence for positive bias. Epidemiology. 2011; 23:1–3.

(a)

(b)

Fig. A.1.

(a) Environment in the first FOCAL evaluation drive (no billboards are present). (b) Environment in the second and third evaluation drive with external distractions (the billboard on the right containing the letter matrix).

Fig. A.2.
Two AMAP screen shots (top panel – forward video view, bottom panel – map view).

Fig. A.3.
University of Massachusetts Amherst driving simulator.

Fig. A.4.
Comparison of the proportion of in-vehicle tasks with maximum glance greater than the threshold value on the x -axis for the FOCAL and Placebo groups.

Fig. A.5.
Comparison of the proportion of glances greater than the threshold value on the x -axis for the FOCAL and Placebo groups for the in-vehicle tasks.

Fig. A.6.
Comparison of the percent of tasks with maximum glance greater than the threshold on the x -axis for the FOCAL and Placebo groups for the out-of-vehicle tasks.

Table A.1

Proportion of in-vehicle tasks in which there was a glance of greater than 2, 2.5, and 3 s.

Measure	Placebo	Focal	Difference
Proportion of tasks in which maximum glance was greater than 2 s	0.756	0.535	0.220 (0.067)
Proportion of tasks in which maximum glance was greater than 2.5 s	0.656	0.346	0.310 (0.075)
Proportion of tasks in which maximum glance was greater than 3 s	0.411	0.206	0.206 (0.079)

Note: Standard errors of the differences are presented in parentheses.

Table A.2

Proportion of glances during in-vehicle tasks that were greater than 2 s, 2.5 s, and 3 s.

Measure	Placebo	FOCAL	Difference
Proportion of glances >2 s	0.308	0.160	0.148 (0.040)
Proportion of glances >2.5 s	0.198	0.095	0.103 (0.033)
Proportion of glances >3 s	0.122	0.044	0.078 (0.024)

Note: Standard errors of the differences are presented in parentheses.

Table A.3

Proportion of in-vehicle tasks in which there was a glance of greater than 2, 2.5, and 3 s (presented separately for map tasks and non-map tasks).

Measure	Placebo	Focal	Difference
<i>Non-map tasks</i>			
Proportion of tasks in which maximum glance was greater than 2 s	0.714	0.471	0.243
Proportion of tasks in which maximum glance was greater than 2.5 s	0.614	0.293	0.321
Proportion of tasks in which maximum glance was greater than 3 s	0.371	0.171	0.200
<i>Map tasks</i>			
Proportion of tasks in which maximum glance was greater than 2 s	0.900	0.750	0.150
Proportion of tasks in which maximum glance was greater than 2.5 s	0.800	0.525	0.275
Proportion of tasks in which maximum glance was greater than 3 s	0.550	0.325	0.225

Table A.4

Proportion of glances during in-vehicle tasks that were greater than 2 s, 2.5 s, and 3 s (presented separately for map tasks and non-map tasks).

Measure	Placebo	Focal	Difference
<i>Non-map tasks</i>			
Proportion of glances >2 s	0.267	0.141	0.127
Proportion of glances >2.5 s	0.178	0.084	0.094
Proportion of glances >3 s	0.111	0.035	0.076
<i>Map tasks</i>			
Proportion of glances >2 s	0.452	0.226	0.226
Proportion of glances >2.5 s	0.268	0.133	0.136
Proportion of glances >3 s	0.162	0.078	0.084

Table A.5

Average total time per task and average number of glances per task.

	Placebo	Focal	Difference	SE	<i>t</i>	<i>P</i>
Avg. total time per task	10.181	8.537	1.644	0.550	2.989	0.005
Avg. no. of glances per task	6.172	6.428	-0.256	0.329	0.777	0.442

Table A.6

Proportion of out-of-vehicle tasks in which there was a glance of greater than 2, 2.5, and 3 s.

Measure	Placebo	Focal	Difference
Proportion of tasks in which maximum glance was greater than 2 s	0.942	0.940	−0.002 (0.018)
Proportion of tasks in which maximum glance was greater than 2.5 s	0.819	0.863	0.044 (0.036)
Proportion of tasks in which maximum glance was greater than 3 s	0.608	0.685	0.078 (0.054)

Note: Standard errors of the differences are presented in parentheses.

Table A.7

Proportion of glances during out-of-vehicle tasks that were greater than 2, 2.5, and 3 s.

Measure	Placebo	FOCAL	Difference
Proportion of glances >2 s	0.883	0.914	0.030 (0.25)
Proportion of glances >2.5 s	0.771	0.836	0.065 (0.041)
Proportion of glances >3 s	0.574	0.665	0.090 (0.056)

Note: Standard errors of the differences are presented in parentheses.