U.S. Department of Transportation
National Highway Traffic Safety Administration

DOT HS 812 379

March 2017

Evaluation of an Updated Version of the Risk Awareness and Perception Training Program for Young Drivers

Disclaimer

Suggested APA Format Citation:

Thomas, F. D., Korbelak, K. T., Divekar, G. U., Blomberg, R. D, Romoser, M. R. E. & Fisher, D. L. (2017, March). *Evaluation of an updated version of the risk awareness and perception training program for young drivers* (Report No. DOT HS 812 379). Washington, DC: National Highway Traffic Safety Administration.

Technical Report Documentation Page

1. Report No. DOT HS 812 379	**2. Government Accession No.**	**3. Recipient Catalog No.**
4 Title and Subtitle Evaluation of an Updated Version of the Risk Awareness and Perception Training Program for Young Drivers		**5 Report Date** March 2017 **6 Performing Organization Code** 258
7. Authors F. Dennis Thomas, III, Kristopher T. Korbelak, Gautam U. Divekar,* Richard D. Blomberg, Matthew R.E. Romoser,* and Donald L. Fisher*		**8 Performing Organization Report No.** 258-1
9 Performing Organization Name and Address Dunlap and Associates, Inc. 110 Lenox Avenue Stamford, CT 06906-2300		**10 Work Unit No. (TRAIS)** **11 Contract or Grant No.** DTNH22-09-D-00138 (Task Order 0005)
12 Sponsoring Agency Name and Address Behavioral Safety Research National Highway Traffic Safety Administration 1200 New Jersey Avenue SE. Washington, DC 20590		**13 Type of Report and Period Covered** Final Report 09/01/10 – 12/31/12 **14 Sponsoring Agency Code**
15 Supplementary Notes *Dr. Divekar, Dr. Romoser, and Dr. Fisher were with the University of Massachusetts at Amherst		

16 Abstract

Previous research suggests newly licensed teen drivers often fail to anticipate where unexpected hazards might materialize. One training program designed to address these apparent deficiencies in knowledge and skills that has shown promise in previous tests is the Risk Awareness and Perception Training (RAPT) program. This project updated RAPT using high definition video and computer simulations to create a more interactive and realistic program. Researchers evaluated the modified program's impact on the behaviors of novice and experienced drivers through the use of a computer-based test and during on-road drives in live traffic on a pre-defined route. Both the novice and experienced driver RAPT-trained groups showed substantial improvement in performance from pre- to post-test with the RAPT trainees hitting almost all of the targets during the computer post-test. The performance differences extended to the eye-tracker data arising from the on-road drives. The RAPT-trained groups hit significantly higher numbers of total primary targets and percentages of targets compared to the control groups. The study also employed a "Think Aloud," or commentary driving, data collection effort. This data collection approach did not reveal any performance differences among the training groups. This study also included a persistence measure using the computer assessment one month after training. Results showed the RAPT-trained groups' target hit rates decreased from the initial post-test to the persistence measure but remained above their baseline hit rates and above the control groups' persistence measure hit rates. Taken together, the results suggest the RAPT revision represented a significant improvement over the previous versions in terms of realism with a similar impact on driver behaviors as measured by a computer assessment and through the use of eye-tracking in a live traffic environment.

17 Key Words Hazard Anticipation Training Teen Driver RAPT Eye-tracker Risk Awareness On-road Study Novice Drivers Driver Education	**18 Distribution Statement** Document is available to the public from the National Technical Information Service www.ntis.gov

19 Security Classification (of this report) Unclassified	**20 Security Classification (of this page)** Unclassified	**21 No. of pages** 42	**22 Price**

Form DOT F 1700.7 (8-72)

ACKNOWLEDGEMENTS

The authors thank the many participants who volunteered for this study. Their willingness to devote their time and return for the persistence measure made the data collection possible. The recruitment of these participants and oversight of the live on-road drives was ably accomplished by Mary Ellen Paciorek of the Pioneer Valley Driving School.

The authors thank the staff of the Arbella Insurance Human Performance Laboratory at the University of Massachusetts at Amherst for their assistance in welcoming and processing participants and for their caring and effective maintenance of the facilities. The authors also acknowledge the excellent technical support from the staff at Applied Science Laboratories, the supplier of the Mobile-Eye tracking equipment used as a primary means of data collection.

Although this study could not have been accomplished without the assistance of these people, the findings and conclusions cited here are those of the authors and are not intended to represent the positions of any of the cooperating agencies or people.

EXECUTIVE SUMMARY

Background and Objectives

Previous research has shown that newly licensed teens often fail to anticipate where unexpected hazards might materialize (Pradhan et al., 2005), and consequently struggle to control the speed, acceleration, and position of their vehicles to avoid hazards (Fisher et al., 2002; Sagberg & Bjørnskau, 2006). A new driver training program designed to address these deficiencies has shown promise: the Risk Awareness and Perception Training (RAPT) program. This project updated RAPT to create a more realistic program and evaluated the new version's impact on novice and experienced drivers' behaviors through a computer-based test and during on-road drives in live traffic.

The objective of this study was to develop and evaluate a hazard anticipation training program. The research questions were:

- *Does participation in RAPT improve hazard detection skills relative to pre-training evaluation and to a comparison (control) group that did not receive the training?*
- *After training, how does hazard anticipation among trainees compare to that of experienced drivers?*
- *To what extent do effects persist one month following training as measured by a computer-based test?*

The modified RAPT program used the same scenarios and presentation order as the prior version but added animations and high definition video. Researchers simplified the instructions and replaced the two-dimensional training diagrams with higher quality graphics. Frame-by-frame photograph sequences from the old version were replaced with high definition video that allowed the user to look right and left 180 degrees. Trainees who failed to click on the hazard after the second time through a training sequence saw a large red oval highlighting the hazard and text explaining why the situation was hazardous. Trainees who responded correctly received a congratulatory message as well as a red oval highlighting the hazard.

Method

Participants. Two hundred five participants enrolled in the study: 103 novice drivers (average age 17) and 102 experienced drivers (average age 24) who had at least one year of driving experience and had driven at least 5,000 miles.

Within each age group, researchers randomly assigned participants to either RAPT or placebo training and then to one of three data collection methods for the on-road portion of the study: eye-tracking, a "Think Aloud" method in which participants described what they saw and thought about as they drove, and both eye-tracking and the Think Aloud technique.

Procedure. Each participant provided consent/assent and completed a demographic questionnaire. A researcher then assigned the participant to a data collection method group and a training group. Participants then completed either RAPT or placebo training. An assessment program embedded in the RAPT and placebo training programs recorded participants' performance.

Following training, participants completed a 2.3-mile on-road drive as directed by a driving instructor. A researcher in the back seat operated the data collection equipment and/or scored the Think Aloud responses. Participants returned for a second session one month after training to repeat the computer-based hazard anticipation assessment.

On-road Drive Scoring. A researcher blind to training condition scored eye-tracker data. The research team defined a hit as a fixation on a target between pre-defined start and end points within a pre-defined radius. For the Think Aloud scoring, researchers identified key words to describe primary targets (e.g., crosswalk or pedestrian); the participant had to say the key words (or a synonym) to score a hit.

Results

Pre- and Post-Testing. The RAPT-trained groups showed significantly greater increases in average number of targets identified in the computer assessment from pre-test to post-test compared to the control groups. A statistically significant two-way interaction of test period and training indicates that, for both experience groups, RAPT training led to a greater increase in targets identified on the post-test than did placebo training.

Analyses of the same measure taken approximately a month after training demonstrated that RAPT training groups obtained lower hit rates as compared to the immediate post-test, but the rates remained above their pre-test performance and above the rates observed for the placebo groups at persistence testing.

On-road Drives. Because the results followed virtually the same pattern, researchers pooled the data from the eye-tracker and eye-tracker plus Think Aloud groups to increase statistical power. Analyses demonstrated a statistically significant training effect in which RAPT participants fixated on more primary targets (average of 19.0 targets) than did placebo-trained participants (average of 15.3 targets). No other main effects or interactions were statistically significant.

Similar to the eye-tracker data, researchers pooled the data from both groups using the Think Aloud technique. Analyses showed no effects of training, experience, or their interaction for the number of primary target key words spoken.

Discussion

RAPT-trained groups showed performance improvements from computer pre- to post-test with the RAPT trainees hitting almost all of the targets during the post-test. Participants retained some of these improvements at one month after training. Effects extended to eye-tracker data collected during on-road drives. RAPT impacted experienced and novice drivers' performance in a similar manner.

Table of Contents

List of Tables

List of Figures

1 BACKGROUND

Analyses of police crash reports indicate that failures of visual scanning (ahead, to the sides, and to the rear), attention maintenance, and speed management contributed, respectively, to 44%, 23% and 21% of the crashes (the causes overlap) among drivers 16 to 19 years old (McKnight & McKnight, 2003). Research has shown that newly licensed teens often fail to anticipate where unexpected hazards might materialize (Pradhan et al., 2005) or control their speed, acceleration and vehicle position to avoid such hazards (Fisher et al., 2002; Sagberg & Bjørnskau, 2006).

A variety of driver training programs strive to address these apparent deficiencies in knowledge and skills. One approach that has shown promise is the Risk Awareness and Perception Training (RAPT) program. This project enhanced a previous version of the RAPT program by including high definition video and computer simulations to create a more interactive and realistic program. Researchers conducted an evaluation of the enhanced program's impact on the behaviors of novice and experienced drivers through the use of a computer-based test and during on-road drives in live traffic.

2 OBJECTIVE

The objective of this study was to develop and evaluate a hazard anticipation training program. A secondary goal was an inexpensive delivery and evaluation method. The primary research questions of interest were:

- *Is participation in the training program associated with improved hazard detection skills relative to pre-training evaluation and a comparison (control) group that did not receive the training?*

- *After training, how does the hazard anticipation skill of trainees compare to experienced drivers?*

- *To what extent do effects of training persist on a computer-based post-test one month after training?*

3 LITERATURE REVIEW

Researchers examined numerous sources for research detailing findings related to training hazard anticipation skills of young drivers. Databases searched included TRIS, PsycInfo, PubMed, Scopus, Academic Search Premier, ABI/Inform Global, NTIS and NHTSA's Behavioral Research Library. After an initial search of the databases, research team members reviewed the abstracts of identified documents. Researchers then read the most relevant in full and included them in the summary of pertinent research presented below.

3.1 Hazard Anticipation and Young Driver Crashes

Motor vehicle crashes are still the largest threat to adolescent health in the United States (Dunlop & Romer, 2010; Shope, 2010). Perhaps in response to this threat, research on teen motor vehicle crashes has increased in recent years. Braitman, Kirley, McCartt, and Chaudhary (2008) used interviews and historical police data to determine the characteristics and factors that contributed to non-fatal teen crashes within the first 8 months of licensure. They found three factors that contributed approximately equally to crashes: failure to detect a hazard (e.g., another vehicle, traffic control device), driving too fast for conditions and losing control of the vehicle. The teenage drivers were at-fault an estimated 75% of the time. Braitman et al. (2008) also found specific crash types with an overrepresentation of teenage drivers, including running off the road, rear ending another vehicle and colliding with another vehicle due to failure to yield right-of-way. Similarly, Chan, Pradhan, Pollatsek, Knodler, and Fisher (2010) suggested that novice drivers commit three primary failures that lead to crashes: failure to anticipate hazards, failure to manage speed and failure to maintain attention.

The primary focus of this research effort is young drivers' hazard anticipation and training programs designed to improve teenage driver hazard anticipation behaviors. As part of one such study, Pradhan et al. (2005) investigated visual search patterns of novice drivers in a driving simulator task. Eye tracking equipment recorded scanning behavior among three comparison groups through 16 simulated drive scenarios that presented risky driving situations. Novice drivers failed to perceive information about potential hazards in many of the scenarios where older drivers detected the hazards.

Huestegge, Skottke, Anders, Müsseler, and Debus (2010) tracked eye movements in relation to hazard perception skills for novice and expert drivers in Germany. Based on previous work, the researchers expected to find differences in visual search patterns between these two groups that would relate to the speed with which they perceived hazards in the driving scenes they were shown. Experts were expected to benefit from top-down processing (knowing where to look) or quicker determination of a spotted object in the visual scene as a hazard. Outcomes of interest were time to the first fixation on a potential hazard and time to decide whether the object was in fact a hazard (participants pressed a key to indicate braking response). Surprisingly, both groups recognized the same number of hazards. Overall, response time was faster for experienced drivers, but the time to first fixation was equivalent across both groups. This suggests that just because a young driver fixates on a hazard does not mean that they recognize the object or situation as hazardous. Studies such as these support the need for hazard anticipation training for young drivers.

3.2 Risk Awareness and Perception Training (RAPT)

The current study is an extension of earlier work using the Risk Awareness and Perception Training (RAPT) program developed by the Arbella Insurance Human Performance Laboratory at the University of Massachusetts Amherst as a means of training hazard anticipation skills to novice drivers. Development of RAPT began in 2005. The first version of RAPT consisted of a PowerPoint slideshow in which participants would see an overview/plan view of various hazardous scenarios. These scenarios typically contained hazards such as hidden pedestrians or moving vehicles. After reading about the scenario, participants used the computer mouse to drag circles to the places in the environment where hazards could be hidden and where they would first be visible to the driver. A later version of RAPT included photos of on-road scenes that contained elements similar to the scenarios depicted in the plan views. Again, participants had to identify hazards and areas where glances should be aimed by dragging circles on to the plan view. In the third version of RAPT, users were given a plan view of scenarios; however, instead of a static photo the participant viewed a progressive slideshow of still photos approaching a hazard from the driver's point of view. Each successive slide simulated the driver getting closer to the hazard. This approach simulated driving, but slowed it down enough for participants to process each scene. Participants could "look" right and left with a mouse click during some slides and were instructed to use their mouse to click on areas where hazards might emerge and in areas they would monitor when approaching the hazard (Fisher, Pollatsek, & Pradhan, 2006).

Driving simulator data used to assess trainee performance following the various versions of RAPT training showed substantial improvements in trainee hazard anticipation skills (Fisher, Pollatsek, & Pradhan, 2006). However, it was important to validate simulator results in the on-road as well. Pradhan, Pollatsek, Knodler, and Fisher (2009) conducted an on-road evaluation of the third version of RAPT in which participants received RAPT training and then drove a 15-mile route in live traffic. Researchers found that trained participants were more likely to fixate on areas of the roadway that contained hazards than were untrained drivers. Trained drivers identified 64% of potential hazards while untrained drivers identified only 37%. Performance was significantly better for RAPT-trained drivers on both near (similar) and far (different) transfer scenarios during the on-road drive.

As part of a large-scale study of RAPT's impact on crashes, Thomas, Blomberg, and Korbelak (2016) modified the RAPT program to streamline its presentation. Their version of RAPT included the same basic concepts as the prior versions, but it employed improved graphics and a more stable delivery platform. Over 2,500 novice drivers completed the RAPT training immediately after receiving their first driver's license that permitted unsupervised driving. These experimental participants also received pre and post (computer-administered) hazard recognition tests. Another 2,500 novice drivers completed the program's pre-test, but received no training or post- test, and therefore constitute a comparison group. The RAPT-trained group showed a large increase in the mean number of hazards identified from pre-test (2.0 out of 9 possible) to post-test (6.8 out of 9 possible). The comparison group showed mean hazard hit scores (1.9 out of 9) similar to the pre-test scores for the RAPT-trained group. In addition, RAPT-trained males showed an approximately 23% reduction in crashes compared to untrained males. Females did not show a similar crash reduction.

3.3 Simulator-Based Training

The cost of the computer processing power and software necessary to run a virtual environment has decreased rapidly in the last 10 years. As a result, the number of simulator-based training approaches has increased. In addition to this increase, the pedagogical strategies employed for the use of driving simulators as a means of training hazard anticipation skills varied considerably across the identified research studies.

In the United Kingdom, Crundall, Andrews, van Loon, and Chapman (2010) evaluated a simulator-based training program that employed commentary driving. Commentary driving involved drivers verbally articulating to the trainer what they were observing, how they were interpreting what they saw and their planned response. It provided the trainer insights into how the drivers scanned the environment, what information they processed, how they processed it and how they intended to react to what they perceived. The researchers developed a simulator-based drive that contained nine hazardous situations of various types. Participants drove this simulation before training as a baseline measure. Training consisted of an on-road drive with an instructor. Participants were provided examples of commentary driving, practiced with a video, and then went on the open road with the instructor and were asked to commentate while they drove. Instructors then provided real-time feedback on missed potential hazards and provided other general safe driving practice instruction. Two weeks later, participants returned to repeat the simulator drive, and received additional feedback. A control group completed the simulated drives without any instruction. Results of a post-test revealed that the trained drivers identified significantly more hazards, had fewer crashes, and reduced their speed sooner in the driving simulator than untrained drivers.

Wang, Zhang, and Salvendy (2010a, 2010b) published two studies of simulator-based hazard anticipation training. In their first study (Wang et al., 2010a), the authors compared hazard anticipation training delivered on a small table-top driving simulator to that of a video-based training program. Both methods involved guided error learning in which participants were allowed to make mistakes and then received feedback on those mistakes. One week later the participants completed a drive in a full-cockpit driving simulator. Participants in the simulator-trained group were better at anticipating hazards in the full-cockpit simulator than participants who completed the video training. However, participants in both groups perceived the training to be highly effective. In their second study (Wang et al., 2010b), the authors evaluated the effectiveness of simulator-based training delivered in the full-cockpit simulator. Participants drove eight scenarios and were provided feedback on their performance and also watched videos of expert performance in the same situations. Scenarios included the fresh green signal, emergency stop, pedestrian, path intrusion and T-intersection turns. The study included a control group that received no training. Participants in the trained group braked earlier for potential hazards and had a lower subjective mental workload than did untrained participants.

A study by Vlakveld et al. (2011) evaluated a new simulator-based version of RAPT called SimRAPT. During the training, participants first drove a 45 to 60 second simulator scenario with a potential hazard that did not actually materialize. The participant reported what they thought the potential hazard was in the drive. They would then drive the scenario again, this time with the hazard materializing. The participants then viewed an interactive feedback video that provided a bird's-eye view of the scenario and the potential hazards. Additionally, two eye-tracking videos showed the trainees examples of drivers performing the drive incorrectly and correctly. Finally, the participant drove the scenario a third time to practice the target skills and strategies communicated in the feedback video. Seven different hazardous scenarios were trained in this fashion. The authors used an eye-tracker to evaluate training transfer on a full cab driving simulator in the lab. In near transfer scenarios, trained drivers fixated on the correct hazard 84% of the time compared to 57% for untrained drivers. In far transfer scenarios, trained drivers fixated on the hazard 71% of the time, compared with 53% for untrained drivers.

3.4 Video, Classroom, and Other Types of Training

Classroom-style lectures and videos are popular education media, primarily because they are relatively easy and inexpensive to produce and do not require complex delivery logistics. They allow for training many students in a relatively short period of time since multiple students can simultaneously view a lecture or video. These types of training strategies typically do not afford the student opportunities to actively practice the skills they are learning. Nonetheless, if carefully designed and targeted, positive training outcomes can be realized relative to hazard anticipation skills.

While not evaluating a particular training program, Borowsky, Shinar and Oron-Gilad (2010) showed substantial age differences in hazard identification skills among drivers who viewed videos of hazardous situations. In their study, participants wore an eye-tracker while viewing videos with various embedded hazards. The study evaluated three populations: young inexperienced, adult experienced and older drivers. The study found that the older adult and experienced driving groups performed equally well at detecting hazards on the videos. No young inexperienced drivers responded to pedestrian hazards, but 40 to 50% of drivers in both older groups did identify the hazards. Younger drivers only responded to salient hazards that represented an immediate threat. These findings strongly suggest that videos could be used to train young drivers and to evaluate their hazard perception skills.

In another study, Carney, McGehee, Lee, Reyes, and Raby (2010) assessed the impact of training that used the output from an onboard vehicle event recorder and weekly parental/teen review of driving performance. Their research used an event triggered video device installed in the car that recorded video of the driver’s face and a second video of the forward roadway. The device recorded g-forces and data such as speed, braking and acceleration from the car’s onboard computer diagnostics interface. The device was programmed to store video and data for 30 seconds before to 30 seconds after any event such as rapid deceleration or acceleration, excessive speed, or high g-forces. Devices recorded drivers’ performance for a one-week baseline, followed by 5 one-week intervention segments and then a final one-week period with no feedback. During the intervention segments, drivers received immediate feedback in the form of a light that illuminated whenever an event was recorded. Teens and parents received weekly event reports and videos for review. The study results showed that events declined by 61%, and driver behavior did not revert to baseline levels during the feedback free period at the end of the study. Drivers with initially high numbers of events benefited most from the training with significant decreases in undesirable events.

McKenna, Horswill, and Alexander (2006) investigated the effects of hazard anticipation training on risk taking behavior in three separate experiments. In the first experiment, training consisted of verbal commentary driving with feedback from an instructor. The study found a decrease in response time for trained drivers when tested using a video-based driving simulation and a decrease in risky driving behavior for novice drivers in the on-road. Experiment two was designed out of concerns that trained drivers may have become hyper-vigilant to hazards during the training and would slow for all situations, regardless of the presence or absence of a hazard. This study found that trained drivers had specific reductions in speed in hazardous situations but not in non-hazardous situations. The third study compared trained police drivers and non-trained police drivers. They found that in general, police had faster reaction times, reduced their speed in advance of hazards, and performed fewer risky driving maneuvers (e.g., following too closely, speeding, gap acceptance) under normal driving conditions than drivers in the prior studies. In addition, those police officers who had taken the advanced training course that included hazard anticipation training showed greater reductions in speeds during the hazardous situations than those officers who had not taken the advanced driving course.

Isler, Starkey, and Williamson (2009) also undertook a study of verbal commentary training of hazard anticipation skills. Participants were trained using video-based verbal commentary driving. Afterwards, they were given a video-based dual perception task. The secondary task involved mimicking the act of steering the car by keeping a cursor inside a moving box on the screen using the mouse. Trained drivers were better than untrained drivers in identifying hazards in the dual-task post-test. The performance of trained novice drivers improved to a level similar to that of adult experienced drivers after training on the post-test and was significantly better than a control group of novice drivers that received no training.

3.5 Computer-Based Hazard Anticipation Testing for Licensure

Using computers for hazard perception testing as part of the requirements for driver licensure is not a new concept. Some jurisdictions in Australia and Europe (e.g., Netherlands, United Kingdom) require computer-based hazard perception tests as part of the driver licensing process. These tests include some form of video or virtual simulation displayed on a computer screen, and require users to "click" on hazards. A publication by Wells et al. (2008) suggested that after the hazard perception test was implemented in the United Kingdom, there was a 3% decrease in young driver crashes during the first year of their driving.

In an effort to create a North American equivalent to the Australian and European tests, Scialfa, et al. (2010) created a hazard perception test for novice drivers with the primary goal of reliably differentiating between novice and experienced drivers. The test displayed a compilation of video-recorded driving scenes on a personal computer. Hazards in the videos included a braking lead vehicle, pedestrian incursion, and construction equipment in the driving lane. The program was designed to be brief and inexpensive enough to be used as part of the testing process for licensure. Participants tapped a 17-inch touch screen to indicate hazards they saw. The computer recorded the spatial coordinates of participant responses and their reaction times to the hazards. Results showed that novice drivers were slower to respond to hazards than experienced drivers (3.2 seconds vs. 2.8 seconds). Despite the somewhat small effects, the authors were still optimistic that the approach was a viable option for licensing agencies.

3.6 Literature Review Summary

Recent research on hazard anticipation training for novice drivers has explored a number of strategies and approaches for training young drivers and for testing their hazard anticipation skills. Many of the studied training strategies showed promise in the laboratory and the on-road. Those that appeared most successful contained some kind of active learning component. The strategies that allowed for a deeper level of processing or practice of the targeted skills appeared most effective at improving hazard anticipation skills not only in the short-term but also in the long term. Also, a number of computer-based tests have been developed to evaluate training programs and as screening tools for licensing agencies. Finally, the extensive use of commentary driving as a training approach suggested that it may be a useful tool for evaluation as well if it could be shown to reliably discriminate different levels of hazard recognition competence. These studies form an excellent background for the current project.

4 UPDATING RAPT

The RAPT program developed during this study used the same scenarios and presentation order as the prior version but with updated animations and high definition video from five synchronized video cameras in a vehicle. To enhance the flow and appearance of the training, the entire package was programmed in Adobe AIR.

To be consistent with the prior versions of the RAPT training, the new program retained the same basic three modules (pre-test, training, post-test) but with numerous enhancements. The new version included the following sections:

1. Welcome screen and participant identification number entry
2. Description of the program and instructions
3. Demonstration drive with the software moving the cursor and showing simulated clicks on screen
4. Practice drive with clicks required in specific locations in order to continue
5. Pre-training assessment drives (9 test drives)
6. Training (RAPT or placebo)
7. Post-training assessment drives (same 9 drives as pre-test)
8. Performance feedback delivered at the end of the post-training assessment drives for RAPT group only
9. “Thank You” screen

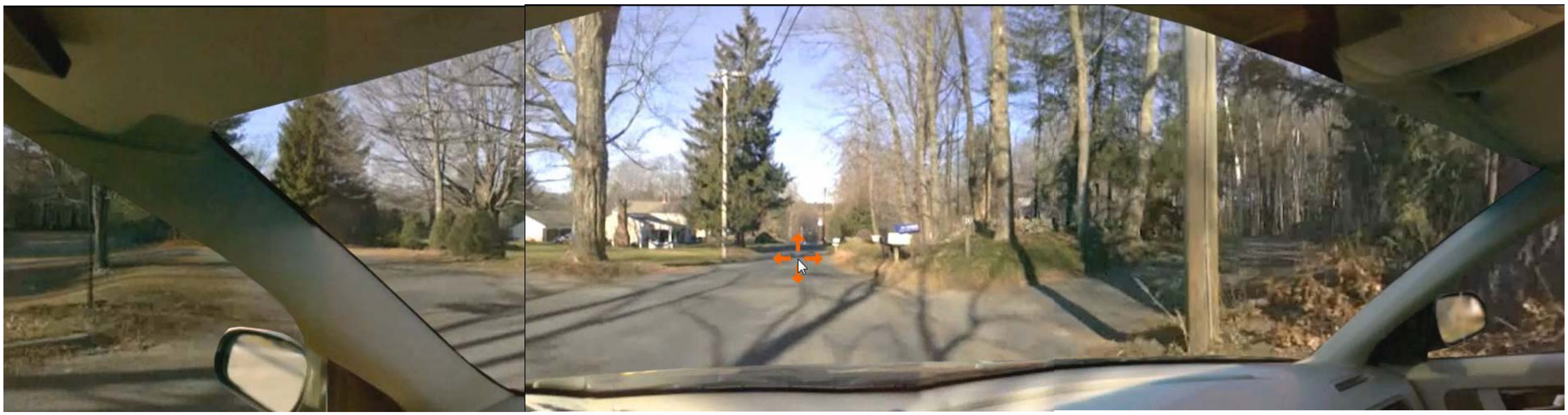
Researchers simplified the instructions to make them easier to read and understand and replaced the two-dimensional training diagrams from the old version with higher quality graphics that included 2- (Figure 1) and 3-dimensional animations (Figure 2). A timer in each training sequence prevented the user from simply clicking through the scenes. The frame-by-frame photograph sequences that represented the drives in the old version were replaced with high definition video that allowed the user to pan right and left 180 degrees as if turning their head while driving (Figure 3).

Just as with the previous version of the program used in the large-scale study in California (Thomas et al., 2016), the number of repetitions of each training sequence was limited to two in order to limit maximum training duration. In the new version, if the trainee did not correctly click on the hazard after the second time through the training sequence, a large red oval highlighted the hazard, and text explained why the situation was hazardous. A trainee who clicked on the hazard correctly received a congratulatory message, and the red oval highlighted the identified hazard (Figure 4) as further reinforcement of the training.

Figure 1. Example of a 2-Dimensional Animation Used in Training.

Figure 2. Example of a 3-Dimensional Animation.

Figure 3. Example of Video Scene



Note. Not all of this Figure is available in the live video scene. The actual view is restricted to approximately one third of the above view and rotates with movements of the mouse as if the user was turning his or her head.

Figure 4. Example of Feedback Scene After Successful Task Completion

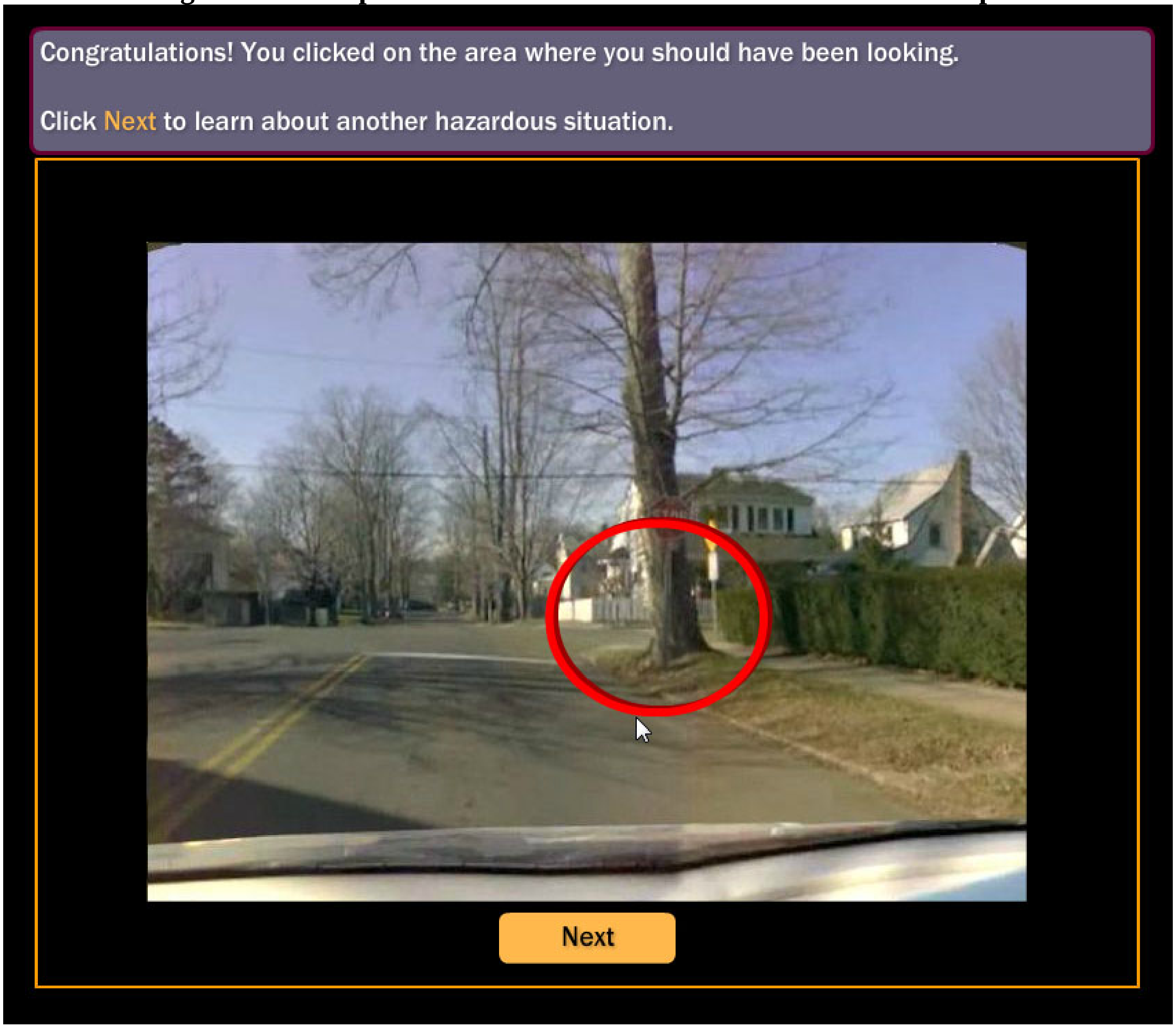
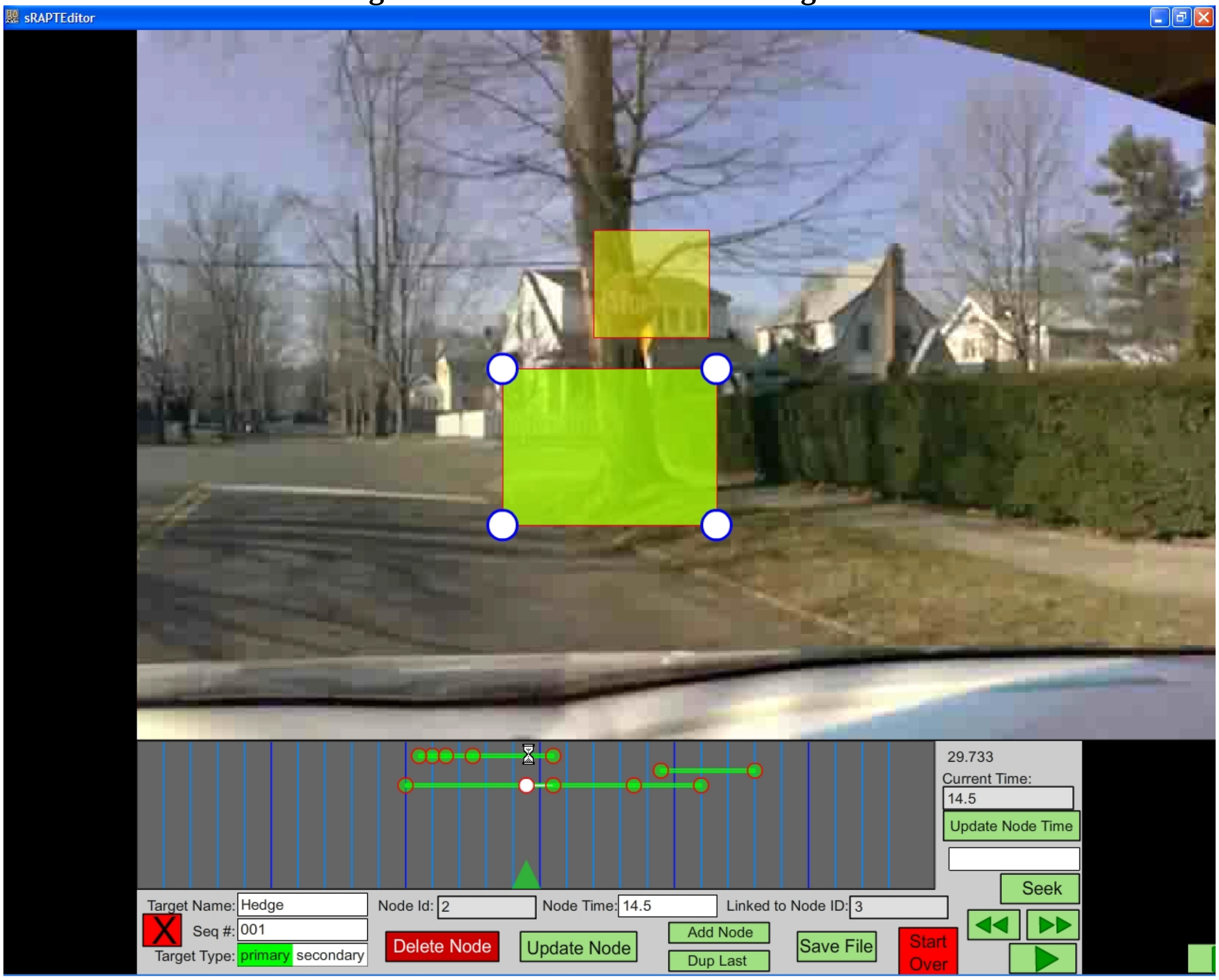


Table 1 provides a brief description of each trained scenario. For each scenario, researchers defined the areas within the video that constituted a "hit" on the primary hazard being trained. Researchers used a computer application to define the target areas in the video as it played (Figure 5). The application extrapolated the target zone between the discrete areas entered by the researchers thereby creating a continuous response surface with appropriate perspective for the particular scene. In addition to the primary hazards, researchers also identified secondary targets. In Figure 5, for example, the hedge on the right was the primary hazard and the Stop Sign (covered by the upper box) was a secondary target. Participants only received feedback concerning clicks on the primary targets, but the RAPT scoring software embedded within the program recorded the number of clicks inside and outside of all of the defined target zones.

Table 1. Description of RAPT Scenarios

Scenario	Description	Primary Hazard/Target Area
1	Two-lane roadway approaching two-way stop; potential hazards obscured by hedge on the right corner.	Hedge on right obscuring crosswalk at intersection
2	Four-lane roadway with mid-block pedestrian crosswalk; potential hazards obscured by cars in left lane.	Area in front of lead vehicle obscuring crosswalk on left
3	Two-lane roadway with hidden drive on left; potential hazard obscured by bushes and trees on the left.	Hidden drive entering from left
4	Two-lane roadway with oncoming vehicle turning left; left-turning vehicle obscures oncoming traffic behind it.	Hidden vehicles behind left-turning vehicle
5	Four-lane roadway with signalized intersection and multiple crosswalks; vehicles in left lane could abruptly change lanes.	Vehicles on left
6	Two-lane roadway in residential neighborhood; lead vehicle turning left must stop for pedestrian on sidewalk.	Pedestrian on sidewalk
7	Three-lane roadway approaching four-way signalized intersection; large vehicle obscures view of signal and oncoming traffic.	Additional traffic light not obscured by large vehicle
8	Two-lane roadway approaching left turn; hill blocks view of oncoming traffic.	Top of hill where oncoming traffic could emerge
9	Two-lane roadway approaching right turn from one-way stop at T-intersection; hill and bend in roadway blocks view of traffic approaching from left.	Top of hill/bend where oncoming traffic could emerge

Figure 5. Tool Used to Define Target Areas



5 METHOD

This study examined the impact of the new version of RAPT on driving behaviors of novice and experienced drivers. The evaluation data included scores on a computer-based test embedded in the RAPT program and on-road drives in actual traffic. Participants returned one month later to take the RAPT computer test again. During the on-road drives, researchers employed two data collection techniques alone and in combination. One technique employed an eye-tracker device to document the location of a participant's gaze while the other involved commentary driving, the "Think Aloud" technique.

5.1 Design

The study used a mixed model design with three between subject variables and multiple dependent variables, one of which was measured on multiple occasions. Between subject variables were the training program completed (RAPT or placebo/control), on-road data collection technique (Think Aloud alone, eye-tracker alone, or Think Aloud and eye-tracker combined), and driver experience/age (Novice 16 to 18-year-olds; Experienced 21 to 30-year-olds). The dependent variables included performance on a computer-based hazard anticipation assessment program (pre-training, immediately after training, persistence 1 month after training), and hazard recognition during the on-road drive that immediately followed training.

5.2 Participants

Researchers aimed to have useable data for at least 15 participants in each experience/age, training, and test condition for a total of 180 participants (2 x 2 x 3). More participants were included in the groups that employed eye-trackers to account for possible data loss related to eye-tracker recording issues that may have occurred (e.g., excessive glare in video). The information below describes the total sample of enrolled participants. The results section provides more details on the actual number of participants in each group with valid data for each data collection method and the persistence measure on the computer one month after training.

A total of 205 participants enrolled in the study; 204 completed the demographics form. Of those, 103 were novice drivers who were 16 to 18 years old (average of 16.6 years, S.D. = 0.64) at the time of the first session, had at least 10 hours of supervised driving experience and had held a learner's permit or Junior Operator License for less than 6 months. The novice group was 61% female and 84% white. The other 101 participants who completed the demographics form (1 experienced driver did not complete the form) were experienced drivers who were 21 to 30 years old (average of 24.2 years, S.D. = 2.52) at the time of the first session, had at least one year of driving experience and had driven at least 5,000 miles. The experienced group was 43% female and 85% white.

Novice drivers were recruited from a local driving school where the driving instructors introduced the study to current and former students who met the eligibility criteria. When a novice driver agreed to participate, a researcher met the novice driver and a parent, explained the study, and gave the parent and novice driver the necessary consent/assent forms before beginning the study. Experienced drivers came from the University campus and surrounding community. Participants received $100 for completing both sessions: $60 after the first session and $40 after the second.

Within each age group, researchers randomly assigned participants to either RAPT or placebo training (the control group). Participants were then randomly assigned to one of three data collection methods for the on-road portion of the study. These data collection groups included the approaches described below:

- Group 1 drivers wore an eye-tracker to collect data on the open road and were simply instructed to drive the route.[1] Researchers later coded the eye tracker data for fixations on potential hazards.

- Group 2 drivers used a “Think Aloud” (commentary driving) technique to describe what they were seeing and their thoughts about the driving situation as they drove. Before they started the drive, participants viewed a three- minute video produced by researchers that described the technique. The drivers were instructed to say what they were looking at and what they were thinking about as they drove. An experimenter in the back seat scored whether the person verbally identification the hazards along the route. If the participant was not using the technique properly, the experimenter prompted him/her to continue describing thoughts as they drove that were related to the drive itself and the driving task.

- Group 3 drivers completed the drive using both the eye-tracker and Think Aloud technique. Again, the experimenter scored whether the person identified the hazards along the route. Researchers later coded the eye-tracker data to document when the person fixated on the hazards. The verbal scores and fixation scores were then compared for this group.

5.3 Materials

Computers. The RAPT and placebo programs operated on four identical Dell Inspiron laptop computers with 15-inch diagonal screens using the Windows operating system.

Hazard Anticipation Assessment Program. Embedded in both the RAPT training and placebo training programs was an assessment program that recorded where in a video a participant clicked using the computer mouse. The assessment program included nine separate simulated drives. The participant was instructed to click on objects or areas in the video where he or she would be looking if actually driving the vehicle. The video allowed the user to pan right and left 180 degrees using the mouse as if the driver’s head was moving. The assessment program recorded whether an individual “hit” (clicked on) specific target areas throughout the drive. Participants completed the program before and after training, and again one month after training. Each testing period included the same nine drives.

---

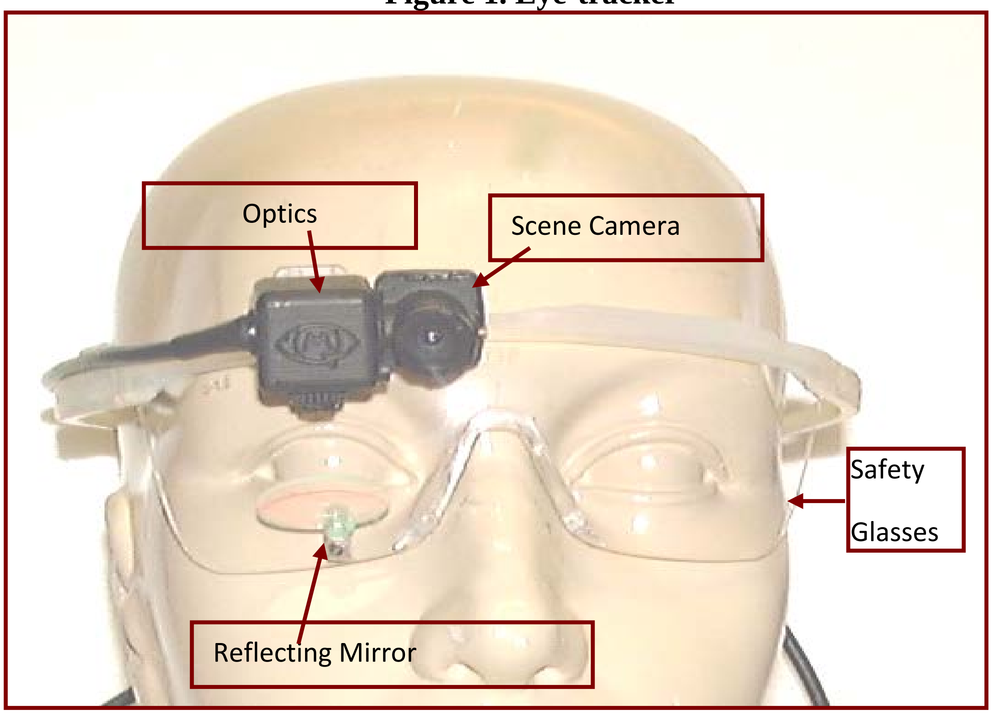
[1] A portable lightweight eye-tracker, the Mobile Eye developed by Applied Science Laboratories, was used to collect the eye-movement data for each driver during the field drives.

Revised RAPT Program. As described above, the RAPT program used for training was the same approach as the prior version but with enhanced animations and high definition video for the training drives. The actual training was embedded between the hazard anticipation assessment program pre-test and post-test.

Placebo Program. The placebo training program included a 10-minute video on automobile maintenance. The video was embedded between the hazard anticipation assessment program pre-test and post-test. The video was chosen because it was related to automobile safety, but it had nothing to do with hazard anticipation or any other topic that could reasonably be expected to impact the scanning behaviors of the participants.

Eye-tracker. The eye-tracker developed by Applied Science Laboratories had a lightweight optical system consisting of an eye camera and a color scene camera mounted on a pair of safety goggles (Figure 1). The system interleaved the images from these two cameras in a single video recording on a laptop computer that was then processed using proprietary ASL software. This software used a standardized calibration file recorded at the start of each session to convert eye movement data to a crosshair representing the driver's point of gaze, superimposed upon the video recorded during the drive. This provided a record of the driver's point of gaze while maneuvering through the live traffic environment. Eye position was sampled at 25 Hz. Head movement was virtually unlimited both vertically and horizontally. The system had an accuracy of 0.58 degrees of visual angle and a resolution of 0.18 degrees of visual angle.

Figure 1. Eye-tracker



Demographics Information. Each participant completed a form containing demographic items such as age, sex, driver license status, and overall driving experience.

Think Aloud Video. A three-minute video showed selected participants how to use the Think Aloud driving technique while on the road. The video was designed to demonstrate the technique without priming participants to focus on specific potential hazards. Project researchers wrote and produced the video that was then installed on each of the laptop computers used for training and testing.

On-Road Test Vehicles. The participating driving school provided the vehicles used in the study. The vehicles were compact cars that had a secondary braking system, which a certified driving instructor could engage if needed. The certified driving instructor also could grab the steering wheel if necessary to redirect the vehicle.

Driving Route. Participants drove a 2.3 mile route through Amherst; drives took place in live traffic. Researchers identified 16 discrete data collection segments along the drive route, some with multiple potential hazards, which resulted in 25 primary target areas being defined and scored. The route included situations similar to those trained by the RAPT program (e.g., hedge blocking view of sidewalk) as well as situations that represented far transfer of training (parked cars near a crosswalk). Researchers chose a route with naturally occurring potential hazards so that that they did not require staging. Therefore, the exact nature of some of the situations varied slightly across participants as a function of prevailing traffic conditions. A data collection drive took 12 to 17 minutes depending on traffic conditions. The Appendix contains a map of the route.

5.4 Procedure: Session 1

In the first session, participants completed the consent/assent forms and the demographic questionnaire. Researchers then randomly assigned participants to either RAPT or placebo training and one of the three on-road data collection methods (eye-tracker, Think Aloud, or both). Each participant then completed the assigned training on the provided laptops.

After training, participants read about the data collection technique to which they had been assigned, watched the Think Aloud video if appropriate, had eye-tracking equipment calibrated if appropriate, and familiarized with the vehicle before beginning the drive. Once the experimenter was satisfied the participant understood the assigned data collection method and study instructions, the researcher turned the participant over to the driving instructor who then directed the driver on the route. The researcher rode in the back seat to operate the data collection equipment and score the Think Aloud as needed. After the drive, researchers debriefed all participants individually and paid them $60 for participation. The session took from 45 to 75 minutes depending on assigned conditions and the speed at which the participant completed the training and the drive.

5.5 Procedure: Session 2

Participants returned to the driving school or UMass one month after training to repeat the computer-based hazard anticipation assessment program. After completing the test, researchers paid participants $40 and briefed them about the study. The second session took approximately 15 minutes.

5.6 Scoring and Analysis

For each of the 25 primary hazard areas in the drive, researchers defined what constituted a glance fixation on the target area in the eye-tracker videos. To be a hit on target, a glance fixation had to take place between the pre-defined start and end points within the scoring segment and be within a pre-defined radius, which expanded as the target came closer. Simply crossing the target zone without fixating did not count. A researcher blind to participant training condition judged whether the glance was indeed an intentional fixation in the target area during the defined scoring period. In many instances, the target area may have been obscured by something such as a parked delivery vehicle or a pedestrian using a crosswalk that would attract the attention of the participant. These situations were noted by the researcher scoring the eye-tracker data and were considered as missing data points.

For the Think Aloud scoring, researchers identified key words that participants could use when describing the primary targets. Participants could use multiple key words to describe some of the 25 primary targets (e.g., crosswalk or pedestrian) that led to the scoring of 34 primary target key words. The participant had to verbalize a primary target key word or a synonym to score a hit for that particular target. Researchers scored another 51 key words participants could mention related to secondary hazards or points of interest (e.g., stop sign). The researcher scored the participant in real time during the drive. All 85 key words can be found in the Think Aloud score sheet in the Appendix.

Analyses focused on assessing whether the RAPT program affected hazard anticipation skills of novice and experienced drivers as compared to groups of similar drivers who completed a placebo training program. A first set of analyses examined data from the computer-based pre-test, post-test, and a persistence measure taken during Session 2, one month after initial training. Mixed model ANOVA was used to test the effect of training over time by experience group.

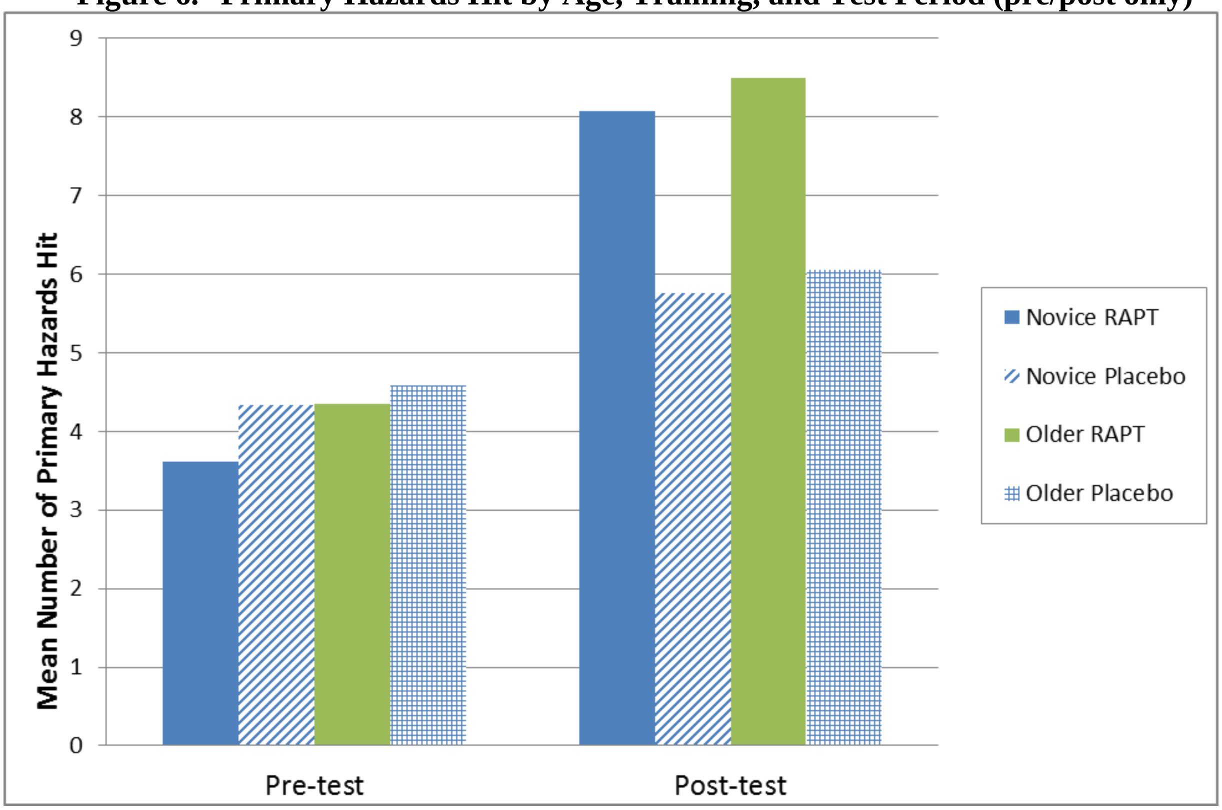
A second set of analyses focused on determining if the on-road test condition affected the eye-tracker and Think Aloud results. Analyses showed no meaningful differences among the eye-tracker, the Think Aloud, and the combined data collection groups, so researchers pooled the data for these three groups for subsequent analyses. Researchers used ANOVA to test for group differences in the mean number of targets hit and percent of valid targets hit.

6 RESULTS

6.1 RAPT Computer Test Pre- and Post-Training

All 205 participants completed the RAPT pre-test and post-test during the first study session. Figure 6 shows the mean number of primary hazards correctly clicked on (out of nine possible) on the pre-test and post-test by age and training group. All experience and training groups obtained similar scores during the pre-test. The RAPT-trained groups, however, showed greater increases in mean number of targets hit from pre-test to post-test compared to the placebo training groups. The mixed model ANOVA showed an overall effect of test period, $F(1, 201) = 370.958$, $p < .001$, which indicates a statistically significant higher mean number of primary targets hit from pre-test to post-test for all groups combined. The two-way interaction of test period and training, $F(1, 201) = 91.021$, $p < .001$, indicates the RAPT training, regardless of participant age group, led to a greater increase in mean number of targets hit on the post-test as compared to placebo training. The two-way interaction of test period and age group and the three-way interaction of test period, age group, and training were not statistically significant at the 0.05 level. The between subjects effect for training was significant, $F(1, 201) = 92.030$, $p < .001$, primarily due to the observed increase for the RAPT trained group on the post-test. The between subjects effect for age group just missed reaching significance, $F(1, 201) = 18.509$, $p = .05$. The between subjects interaction of training and age group was nonsignificant ($p > .05$).

Figure 6. Primary Hazards Hit by Age, Training, and Test Period (pre/post only)



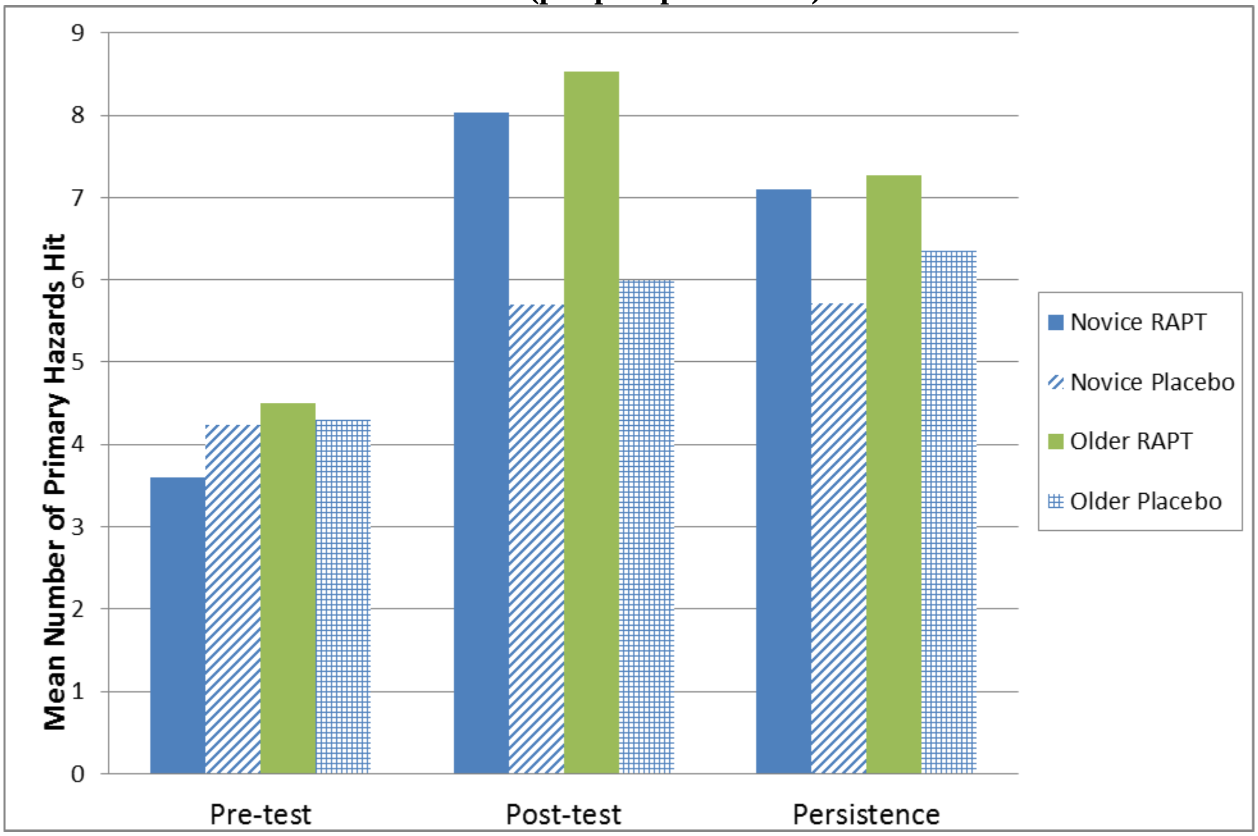
Sample size: Novice RAPT = 52, Novice Placebo = 51, Older RAPT = 51, Older Placebo = 51

6.2 RAPT Computer Test Persistence Measure

Analyses of data from the computer persistence measure taken approximately a month after the initial training only included participants who completed ***both*** study sessions (N = 177). The mixed ANOVA results showed a test period main effect and interaction of test period and training. Figure 7 shows the RAPT-trained groups tended to score lower during the persistence measure than during the post-test immediately after training, but their persistence scores remained higher than the placebo training groups' persistence scores. Given the observed quadratic trends, researchers conducted separate analyses comparing the persistence measure to the post-test and then comparing the persistence measure to the pre-test to provide a more easily understood analysis of the data.

When comparing the persistence data to the post-test data, there was an overall effect of test period, $F(1, 173) = 10.877$, $p = .001$ indicating a statistically significant reduction in target hits across all groups. However, the two-way interaction of test period and training was also significant, $F(1, 173) = 21.571$, $p < .001$, with the RAPT group showing a lower hit rate at the persistence measure than at the post-test while the placebo groups actually increased their mean number of hits slightly from post-test to persistence measure. The two-way interaction of test period and age group and the three-way interaction of test period, age group, and training were not significant at the 0.05 level. The between subjects effect for training was significant, $F(1, 173) = 63.132$, $p < .001$. The between subjects interaction of training and age group was not statistically significant ($p > .05$).

Figure 7. Primary Hazards Hit by Age, Training, and Test Period (pre/post/persistence)



Ns: Novice RAPT = 50 Novice Placebo = 47 Older RAPT = 40 Older Placebo = 40

Though the persistence scores for the RAPT groups decreased from the post-test period, it was important to determine if the persistence scores were significantly higher than the pre-test measure. Similar to the prior analyses, there was an overall effect of test period, $F(1, 173) = 201.255$, $p < .001$indicating higher overall mean number of targets hit from pre-test to persistence measure for all groups combined. The two-way interaction of test period and training was also significant, $F(1, 173) = 15.636$, $p < .001$, indicating that while all groups showed higher mean number of targets hit in the persistence measure, the increase for the RAPT groups was greater than for the placebo groups. The two-way interaction of test period and age group and the three-way interaction of test period, age group, and training were not statistically significant ($p > .05$).

6.3 On-Road Drives: Eye-Tracker Data

The data collection approach involved the use of an eye-tracker with two separate groups of participants. One group involved use of only the eye-tracker data while the other included eye-tracker and Think Aloud data collection techniques. Researchers examined the data from each data collection technique both to determine if the RAPT training produced improved performance and to see if the eye-tracker and Think Aloud approaches yielded equivalent results.

For the data coded from the eye-tracker videos, researchers compared hit and miss statistics for each of the 25 primary targets across participants in the two data collection groups. Overall, three of the 25 targets showed minor differences in hit rates by data collection technique. Researchers also compared rates of missing data across groups and found no differences among the groups. Because their results followed virtually the same pattern, researchers pooled eye-tracker data from both groups to increase statistical power.

Researchers used the GLM univariate ANOVA procedure to investigate differences in the total number of targets hit by training and age group. The first analysis examined mean differences in the total number of targets hit out of a maximum of 25 primary targets. There was a significant training effect, F (1, 119) = 37.30, p < .001, indicating that on average, RAPT participants fixated on more primary targets (M = 18.97, SD = 3.22) than did placebo-trained participants (M = 15.31, SD = 3.42). No other main effects or interactions were statistically significant. Table 2 provides means for all ages and training groups.

Table 2. Eye-Tracker: Mean Number of Primary Targets Hit

Training Group	Age Group	(N)	Mean	(SD)
RAPT	Novice	(28)	18.96	(3.21)
	Older	(30)	18.97	(3.27)
	Total	(58)	18.97	(3.22)
Placebo	Novice	(33)	14.55	(3.44)
	Older	(32)	16.09	(3.27)
	Total	(65)	15.31	(3.42)
Total	Novice	(61)	16.57	(3.99)
	Older	(62)	17.48	(3.55)
	Total	(123)	17.03	(3.79)

The second analysis examined the mean percentage of total targets hit by group. A percentage of targets hit was calculated for each individual and then averaged across all participants in a group to provide a group mean percentage of targets hit. The denominator used to calculate an individual's percentage of targets hit could be different than that for other participants because of differences in the amount of missing data (e.g., targets blocked as described above). Analysis revealed a significant training effect, F(1, 119) = 53.97, p < .001 indicating that, on average, RAPT participants had a greater percentage of targets hit (M = 83.34, SD = 9.99) than placebo-trained participants (M = 68.20, SD = 13.05). There was also a significant age group effect, F (1, 119) = 6.38, p = .013 with older participants hitting a greater percentage of targets (M = 78.19, SD = 11.05) than younger participants (M = 72.44, SD = 15.88). There was no significant interaction between training and age group. Table 3 provides means for all age and training groups. The percentage of each age and training group hitting the individual targets can be found in the Appendix.

Table 3. Eye-Tracker: Mean Percentage of Primary Targets Hit for Valid Data

Training Group	Age Group	(N)	Mean%	(SD)
RAPT	Novice	(28)	82.57%	(11.76)
	Older	(30)	84.05%	(8.13)
	Total	(58)	83.34%	(9.99)
Placebo	Novice	(33)	63.84%	(13.79)
	Older	(32)	72.71%	(10.67)
	Total	(65)	68.20%	(13.05)
Total	Novice	(61)	72.44%	(15.88)
	Older	(62)	78.19%	(11.05)
	Total	(123)	75.34%	(13.91)

6.4 On-Road Drives: Think Aloud Data

Although there were only 25 primary targets defined during the drive and scored for the eye-tracker data, some targets could have been described using multiple words. As such, researchers defined 34 key words associated with primary targets along the drive. Another 51 secondary key words were included to represent secondary hazards. A copy of the score sheet is in the Appendix. Researchers compared hit and miss statistics (count and percentages) for participants who wore eye-trackers and also provided Think Aloud data and participants who only provided Think Aloud data. These analyses showed no meaningful differences in key word hits by data collection condition. Since no differences were found, researchers pooled the data from the data collection techniques to provide more statistical power.

The GLM univariate ANOVA procedure was then used to investigate differences in the total number of primary and secondary target key words across all scenarios by training and age group. All 85 key words from the 16 segments were included in this analysis. The dependent variable in this analysis was the total number of key words spoken by participants during the drive. There were no effects of training, F (1, 125) = .00, p = .997, age group, F (1, 125) = 0.827, p = .365, or their interaction, F (1, 125) = .125, p = .725. Table 4 provides means for all ages and training groups.

Table 4. Think Aloud: Mean Number Key Words Out of 85 Possible

Training Group	Age Group	(N)	Mean	(SD)
RAPT	Novice	(34)	34.65	(14.14)
	Older	(33)	33.21	(13.12)
	Total	(67)	33.94	(13.56)
Placebo	Novice	(31)	35.55	(15.33)
	Older	(31)	32.29	(15.96)
	Total	(62)	33.92	(15.61)
Total	Novice	(65)	35.08	(14.61)
	Older	(64)	32.77	(14.46)
	Total	(129)	33.93	(14.52)

The GLM univariate ANOVA procedure was then used to investigate differences in the total number of primary target key words spoken. There were no effects of training, F (1, 125) = .055, p = .814, age group, F (1, 125) = .21, p = .648, or their interaction, F (1, 125) = .565, p = .453. Table 5 provides means for all ages and training groups. Since there was no missing data using this data collection technique, there was no need to conduct an analysis of the percentage of key words spoken. For descriptive purposes, however, the percentage of each group speaking each of the 34 key words can be found in the Appendix.

Table 5. Think Aloud: Mean Number Primary Target Key Words Out of 34 Possible

Training Group	Age Group	(N)	Mean	(SD)
RAPT	Novice	(34)	15.12	(6.19)
	Older	(33)	15.45	(5.64)
	Total	(67)	15.28	(5.88)
Placebo	Novice	(31)	15.71	(7.18)
	Older	(31)	14.32	(6.97)
	Total	(62)	15.02	(7.05)
Total	Novice	(65)	15.40	(6.64)
	Older	(64)	14.91	(6.30)
	Total	(129)	15.16	(6.45)

7 SUMMARY AND DISCUSSION

The version of the RAPT program developed for this project included state-of-the-art graphics and live-action video to increase the realism of the training while maintaining the fundamental themes of the prior versions. The study results showed the new version increased hazard anticipation skills as measured by the computer assessment test completed before and after training. The RAPT-trained groups showed substantial increases in performance from pre- to post-test with the RAPT trainees hitting almost all of the targets during the post-test. The placebo groups showed small increases in performance from pre- to post-test that suggests the assessment program itself may represent some form of hazard anticipation training, especially because it repeats the same scenarios. The increased performance after training is consistent with past studies of the RAPT approach, and the higher hit rates during the pre-test measure in this study may be due to the increased realism and ease of use of the new assessment program's 180-degree videos compared to the still frame approach used in earlier versions.

The performance differences extended to the eye-tracker data collected during the on-road drives. The RAPT-trained groups showed significantly higher numbers of total primary targets hit and percentages of targets hit compared to the control groups. RAPT appeared to affect experienced older and novice younger drivers in a similar manner with both groups showing improved performance on the on-road drives after completing the training. In fact, the novice RAPT-trained drivers increased their performance almost to the level observed for the experienced drivers. These results are encouraging since most of the on-road drive scenarios varied somewhat from those used in training.

This study included the collection of data using the Think Aloud (or commentary driving) technique in which drivers were asked to verbalize what they were thinking about and looking at as they drove. Including the Think Aloud technique permitted an analysis of whether it is a suitable surrogate for the more complex and expensive use of an eye-tracker. Analyses of the data collected using the Think Aloud approach did not show any differences in key word hit rates among the training and age groups. In fact, additional analyses not presented in the results section showed very little relationship between eye-tracker hits/misses and key word hits/misses for the group that had data collected using both approaches. These findings suggest that the Think Aloud technique may have limited use in studies such as the one conducted here where brief glances at specific targets in a live traffic environment are of interest. The commentary driving technique has been beneficial in previous studies in which it is focused on more macroscopic issues such as whether the driver detects a pedestrian or bicyclist, and nothing found here invalidates those earlier findings.

This study also included the addition of a persistence measure taken one month after initial training in which participants returned to complete the computer assessment program again. Results showed the RAPT-trained groups' target hit rates decreased from the post-test to the persistence measure but remained above their pre-test hit rates and above the control groups' persistence measure hit rates. The control group's hit rates remained relatively stable from the post-test to persistence measure. It is not surprising that hit rates would go down for the RAPT-trained group since they achieved such high levels of performance on the initial post-test and that test occurred immediately after seeing the scenarios twice—once in the pre-test and once in the training. These results do suggest, however, that advanced skills such as hazard anticipation might benefit from booster training or reminders aimed at keeping performance levels high until drivers have developed ingrained habit patterns.

As with all research, the study had limitations. This research effort involved recruiting volunteer participants from a single driving school and from a University campus. While strong, the results may not be generalizable to a more diverse, randomly selected population that may be less interested in completing driver safety training. Future research should consider testing the impacts of the training on a group of drivers from a more diverse background and in a State with different graduated driver licensing laws.

The maintenance of participant safety and the collection of data during the on-road drives necessitated having a driving instructor sitting in the right front passenger seat and a researcher sitting in the back seat. The presence of these two individuals could have affected the driving and/or glance behaviors of the participants. Future research may wish to examine glance behaviors of RAPT-trained drivers using more naturalistic observation techniques.

The control groups improved their performance on the computer-based assessment test that suggests the assessment program itself may have been a form of training. Simply completing this assessment may have led to changes in driving behaviors if the participants were trying to determine what the study was about. If anything, this tended to reduce the magnitude of any observed differences. Future research may wish to forego giving the computer-based assessment program to the comparison group to see if the observed differences on the on-road drives become greater than those found in this study.

Taken everything together, the results suggest that the RAPT revision used in this project represented a significant improvement over the previous versions in terms of realism and had a similar impact on driver behaviors as measured by a computer assessment and use of eye-tracking in a live traffic environment.

8 REFERENCES

Borowsky, A., Shinar, D., & Oron-Gilad, T. (2010). Age, skill, and hazard perception in driving. *Accident Analysis and Prevention, 42*(4), 1240-1249. doi: 10.1016/j.aap.2010.02.001

Braitman, K. A., Kirley, B. B., McCartt, A. T., & Chaudhary, N. K. (2008). Crashes of novice teenage drivers: Characteristics and contributing factors. *Journal of Safety Research, 39*(1), 47-54.

Carney, C., McGehee, D. V., Lee, J. D., Reyes, M. L., & Raby, M. (2010). Using an event-triggered video intervention system to expand the supervised learning of newly licensed adolescent drivers. *American Journal of Public Health, 100*(6), 1101-1106. doi: 10.2105/ajph.2009.165829

Chan, E., Pradhan, A. K., Pollatsek, A., Knodler, M. A., & Fisher, D. L. (2010). 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, 13*(5), 343-353. doi: 10.1016/j.trf.2010.04.001

Crundall, D., Andrews, B., van Loon, E., & Chapman, P. (2010). Commentary training improves responsiveness to hazards in a driving simulator. *Accident Analysis and Prevention, 42*(6), 2117-2124. doi: 10.1016/j.aap.2010.07.001

Dunlop, S. M., & Romer, D. (2010). Adolescent and young adult crash risk: Sensation seeking, substance use propensity and substance use behaviors. *Journal of Adolescent Health, 46*(1), 90-92.

Fisher, D. L., Laurie, N. E., Glaser, R., Connerney, K., Pollatsek, A., Duffy, S. A., & Brock, J. (2002). The use of an advanced driving simulator to evaluate the effects of training and experience on drivers’ behavior in risky traffic scenarios. *Human Factors, 44*, pp. 287-302.

Fisher, D. L., Pollatsek, A. P., & Pradhan, A. (2006). Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Injury Prevention, 12,* 25-29. doi: 10.1136/ip.2006.012021

Huestegge, L., Skottke, E.-M., Anders, S., Müsseler, J., & Debus, G. (2010). The development of hazard perception: Dissociation of visual orientation and hazard processing. *Transportation Research Part F: Traffic Psychology and Behaviour, 13*(1), 1-8.

Isler, R. B., Starkey, N. J., & Williamson, A. R. (2009). Video-based road commentary training improves hazard perception of young drivers in a dual task. *Accident Analysis and Prevention, 41*(3), 445-452. doi: 10.1016/j.aap.2008.12.016

McKenna, F. P., Horswill, M. S., & Alexander, J. L. (2006). Does anticipation training affect drivers' risk taking? *Journal of Experimental Psychology-Applied, 12*(1), 1-10. doi: 10.1037/1076-898x.12.1.1

McKnight, A. J., & McKnight, S. A. (2003). Young novice drivers: Careless or clueless. *Accident Analysis and Prevention, 35*, 921-925.

Pradhan, A. K., Hammel, K. R., DeRamus, R., Pollatsek, A., Noyce, D. A., & Fisher, D. L. (2005). Using eye movements to evaluate effects of driver age on risk perception in a driving simulator. *Human Factors, 47*(4), 840-852.

Pradhan, A. K., Pollatsek, A., Knodler, M., & Fisher, D. L. (2009). 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, 52*(6), 657-673. doi: 10.1080/00140130802550232

Sagberg, F., & Bjørnskau, T. (2006). Hazard perception and driving experience among novice drivers., *Accident Analysis and Prevention,3,* 407-414.

Scialfa, C. T., Deschênes, M. C., Ference, J., Boone, J., Horswill, M. S., & Wetton, M. (2010). A hazard perception test for novice drivers. *Accident Analysis and Prevention, 43,* 204-208.

Shope, J. T. (2010). Adolescent motor vehicle crash risk: What's needed to understand and reduce the risk? *Journal of Adolescent Health, 46*(1), 1-2.

Thomas, F.D., Blomberg, R.D., & Korbelak, K.T. (2016). *Evaluation of the safety benefits of the risk awareness and perception training program for novice teen drivers (Report No. DOT HS 812 235).*. Washington DC: National Highway Traffic Safety Adminstration.

Vlakveld, W., Romoser, M. R. E., Mehranian, H., Diete, F., Pollatsek, A., & Fisher, D. L. (2011). Does the experience of crashes and near crashes in a simulator-based training program enhance novice drivers' visual search for latent hazards? *Transportation Research Record, 2265,* 153–160.

Wang, Y. B., Zhang, W., & Salvendy, G. (2010a). A comparative study of two hazard handling training methods for novice drivers. *Traffic Injury Prevention, 11*(5), 483-491. doi: 10.1080/15389588.2010.489242

Wang, Y. B., Zhang, W., & Salvendy, G. (2010b). Effects of a simulation-based training intervention on novice drivers' hazard handling performance. *Traffic Injury Prevention, 11*(1), 16-24. doi: 10.1080/15389580903390631

Wells, P., Tong, S., Sexton, B., Grayson, G., & Jones, E. (2008). *Cohort II: A study of learner and new drivers, Volume 1: main report* (Road Safety Research Report No. 81) London: Department for Transport, London.

Appendix

Eye-tracker: Percentage of Group Hitting a Primary Target

Target	Novice RAPT	Novice Control	Experienced RAPT	Experienced Control
Scenario 1 - Traffic Light	100.0%	100.0%	96.7%	100.0%
Scenario 1 - Right Side of Crosswalk	84.0%	50.0%	78.6%	81.5%
Scenario 1 - Left Side of Crosswalk	66.7%	36.0%	65.4%	36.0%
Scenario 2 - Ahead of Bush on Left	59.3%	36.4%	63.3%	53.3%
Scenario 2 - Reveal Point on Left Side	92.9%	93.8%	96.6%	92.9%
Scenario 3 - Reveal Point on Curve	100.0%	87.9%	100.0%	96.6%
Scenario 3 - Left Side of Crosswalk	72.0%	46.9%	79.3%	59.3%
Scenario 4 - Ahead of Right Bush	100.0%	87.5%	100.0%	100.0%
Scenario 5 - Adjacent Road on Approach	60.7%	33.3%	62.1%	43.8%
Scenario 6 - Sidewalk on Right	80.8%	50.0%[a]	89.7%	72.4%
Scenario 6 - Extreme Left	92.6%	81.8%	100.0%	93.1%
Scenario 7 - Sidewalk on Right	88.5%	71.0%	89.3%	73.3%
Scenario 8 - Sidewalk on Right	90.9%	52.0%	84.2%	75.0%
Scenario 9 - Sidewalk on Right	100.0%	80.0%	95.2%	95.5%
Scenario 10 - Opposing Traffic	100.0%	100.0%	100.0%	100.0%
Scenario 10 - Right Side of Crosswalk	96.2%	75.0%	96.6%	92.9%
Scenario 10 - Left Side of Crosswalk	72.7%	42.3%	85.7%	52.2%
Scenario 11 - Ahead of Right Hedge	81.5%	54.5%	55.2%	34.4%
Scenario 12 - Driveway on Right	39.1%	26.7%	46.2%	37.9%
Scenario 13 - Right Side of Crosswalk	57.7%	32.3%	75.0%	58.1%
Scenario 14 - Parked Cars on Right	100.0%	84.8%	100.0%	100.0%
Scenario 14 - Ahead of Parked Cars Left	100.0%	87.5%	100.0%	100.0%
Scenario 15 - Right Side of Crosswalk	65.4%	55.2%	70.8%	59.3%
Scenario 16 - Right Side of Crosswalk	76.9%	68.8%	95.8%	89.7%
Scenario 16 - Left Side of Crosswalk	77.3%	50.0%	56.5%	22.2%

Think Aloud: Percentage of Group Verbalizing a Primary Key Word

Key Word	Novice RAPT	Novice Control	Experienced RAPT	Experienced Control
Scenario 1 - Crosswalk	32.4%	48.4%	60.6%	54.8%
Scenario 1 - Pedestrian	47.1%	58.1%	69.7%	54.8%
Scenario 1 - Traffic Light	94.1%	93.5%	84.8%	83.9%
Scenario 2 - Hedge/Bush	41.2%	41.9%	66.7%	38.7%
Scenario 2 - Clear View	76.5%	77.4%	54.5%	51.6%
Scenario 3 - Curve Blocking	47.1%	45.2%	63.6%	51.6%
Scenario 3 - Crosswalk	5.9%	29.0%	9.1%	22.6%
Scenario 3 - Pedestrian	17.6%	32.3%	12.1%	16.1%
Scenario 4 - Hedge/Trees	23.5%	32.3%	36.4%	29.0%
Scenario 5 - Merging Road	29.4%	29.0%	36.4%	48.4%
Scenario 6 - Crosswalk	44.1%	38.7%	39.4%	35.5%
Scenario 6 - Pedestrian	61.8%	58.1%	51.5%	38.7%
Scenario 7 - Crosswalk	64.7%	71.0%	72.7%	74.2%
Scenario 7 - Pedestrian Crossing Sign	20.6%	19.4%	3.0%	0.0%
Scenario 7 - Pedestrian	73.5%	71.0%	57.6%	67.7%
Scenario 8 - Crosswalk	64.7%	64.5%	87.9%	80.6%
Scenario 8 - Pedestrian	52.9%	61.3%	69.7%	67.7%
Scenario 9 - Crosswalk	58.8%	67.7%	72.7%	64.5%
Scenario 9 - Pedestrian Crossing Sign	23.5%	9.7%	3.0%	0.0%
Scenario 9 - Pedestrian	70.6%	61.3%	69.7%	80.6%
Scenario 10 - Oncoming Traffic	91.2%	90.3%	84.8%	90.3%
Scenario 10 - Crosswalk	26.5%	38.7%	39.4%	25.8%
Scenario 10 - Look Left	38.2%	41.9%	39.4%	41.9%
Scenario 11 - Hedge	47.1%	45.2%	54.5%	32.3%
Scenario 12 - Hidden Drive	11.8%	25.8%	12.1%	12.9%
Scenario 13 - Crosswalk	32.4%	41.9%	36.4%	32.3%
Scenario 13 - Pedestrian	47.1%	41.9%	51.5%	45.2%
Scenario 14 - Cars Blocking	29.4%	32.3%	30.3%	12.9%
Scenario 14 - Left-Right-Left	82.4%	74.2%	81.8%	83.9%
Scenario 15 - Crosswalk	55.9%	48.4%	39.4%	35.5%
Scenario 15 - Pedestrian Crossing Sign	8.8%	6.5%	0.0%	3.2%
Scenario 15 - Pedestrian	50.0%	32.3%	21.2%	25.8%
Scenario 16 - Pedestrian	20.6%	16.1%	18.2%	12.9%
Scenario 16 - Crosswalk	20.6%	25.8%	15.2%	16.1%

Think Aloud Score Sheet

	Scenario	Key Words	Score 1-4
1	**Left Across Traffic** North Pleasant with left turn onto Kellogg Ave	crosswalk pedestrian oncoming traffic light bus stop parked cars ped signal	
2	**Right Turn With Left Reveal** Kellogg with right turn onto Triangle	stop sign stop line creeping hedge/bush clear view	
3	**Left Turn With Right Reveal** Triangle with left turn onto Matoon	curve blocking look left crosswalk pedestrian oncoming cars trees/bushes obscured	
4	**Left Across Traffic Trees Blocking View** Matoon with left turn onto Triangle	stop sign multiple looks hedge/trees creeping obscured	
5	**Merging Road** Lessey with Tyler entering from right	merging road trees/bushes obscured	
6	**Right Turn at 2-way Stop** Lessey with right turn onto Main Street	stop sign trees/bushes parked cars crosswalk pedestrians oncoming traffic creeping	
7	**Midblock Crosswalk 1** Straight on Main just before Boltwood Walk	parked cars crosswalk ped crossing sign oncoming traffic pedestrian cross traffic obscured	
8	**Midblock Crosswalk 2** Straight on Main just before Boltwood Ave.	parked cars crosswalk oncoming traffic pedestrian obscured	

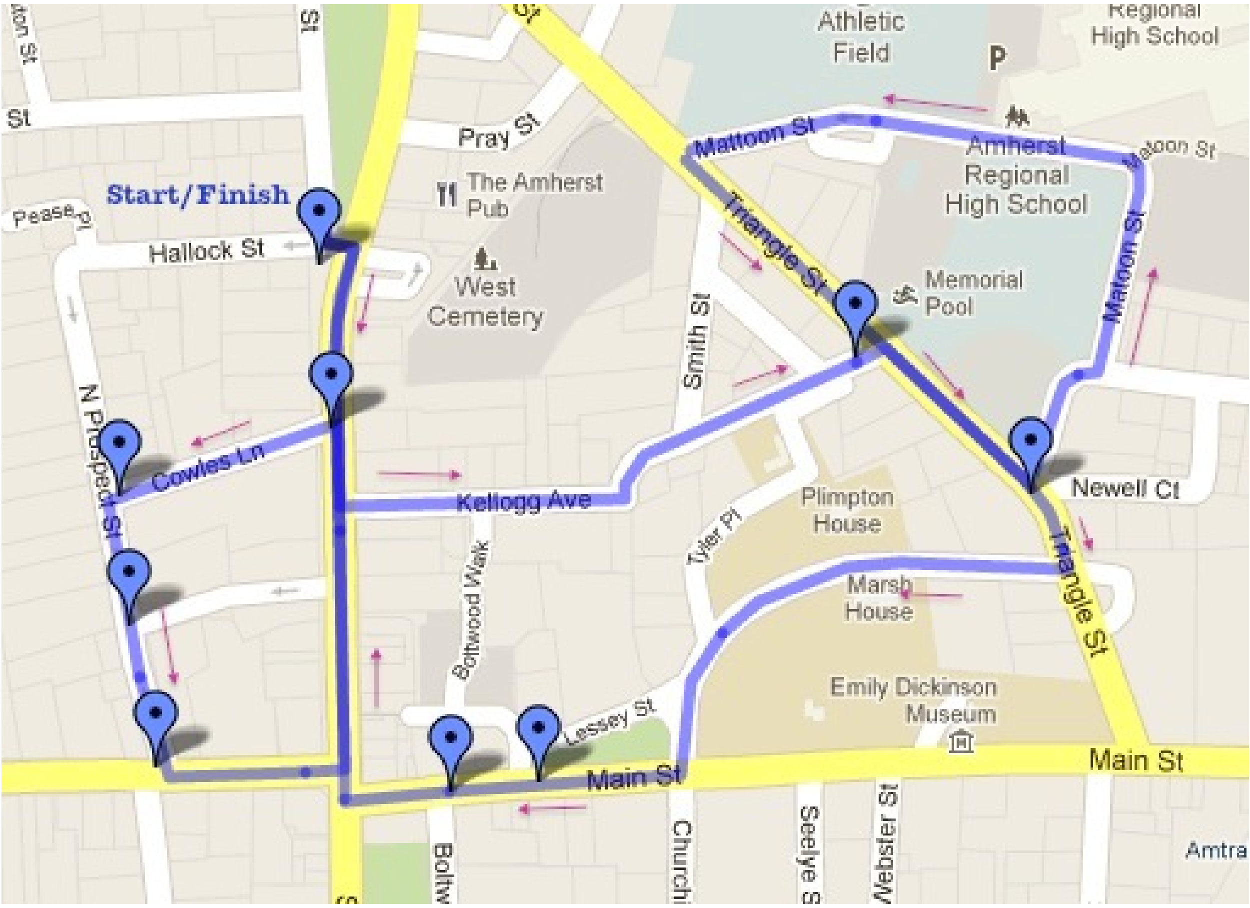
Think Aloud Score Sheet

	Scenario	Key Words	Score 1-4
9	**Midblock Crosswalk 3 (Starbucks)** Straight on North Pleasant	parked cars crosswalk ped crossing sign oncoming traffic pedestrian obscured	
10	**Left Across Traffic** North Pleasant with left onto Cowles Lane	oncoming traffic crosswalk look left	
11	**Hedge Blocking View** Cowles with left onto North Prospect	stop sign hedge parked cars pedestrian creeping obscured	
12	**Hidden Drive on Left** Straight on North Prospect	hidden drive parking sign trees/bushes obscured	
13	**Crosswalk at Stop Sign** North Prospect approaching Amity	stop sign crosswalk creeping pedestrians	
14	**Left Across Traffic at 2-Way Stop** North Prospect with left onto Amity Street	creeping cars blocking trees/bushes left-right-left oncoming traffic creeping	
15	**Midblock Crosswalk 5** North Pleasant second midblock crosswalk	parked cars crosswalk ped crossing sign oncoming traffic pedestrian obscured	
16	**Left Across Traffic** North Pleasant with left onto Hallock Street	oncoming traffic look left pedestrian crosswalk	

How well did the participant apply the Think Aloud technique? 1 2 3 4 5 6 7 8 9 10
Very Poorly ... Very Well

How well did the person drive the course? 1 2 3 4 5 6 7 8 9 10
Very Poorly ... Very Well

On-road Drive Route in Amherst, MA



DOT HS 812 379
March 2017

U.S. Department
of Transportation

National Highway
Traffic Safety
Administration

NHTSA

12913-031317-v3b