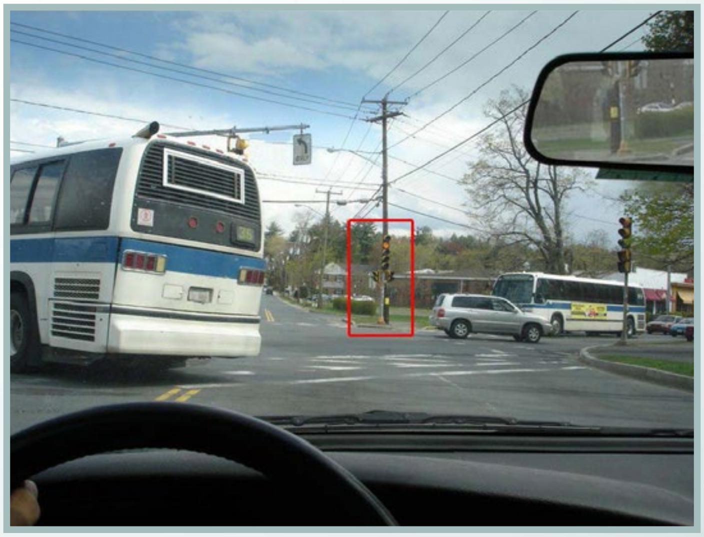
Evaluation of the Safety Benefits of the

Risk Awareness and Perception Training Program for Novice Teen Drivers



January 2016

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

NHTSA
www.nhtsa.gov

DISCLAIMER

Suggested APA Format Citation:

Thomas, F. D., Rilea, S. L., Blomberg, R. D., Peck. R. C., & Korbelak, K. T. (2016, January). *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 Administration.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT HS 812 235		

4. Title and Subject

Evaluation of the Safety Benefits of the Risk Awareness and Perception Training Program for Novice Teen Drivers

5. Report Date
January 2016

6. Performing Organization Code
280

7. Authors

F. Dennis Thomas, Stacy Rilea, Richard D. Blomberg, Raymond C. Peck, & Kristopher T. Korbelak

8. Performing Organization Report No.

280-1

9. Performing Organization Name and Address

Dunlap and Associates, Inc.
110 Lenox Avenue
Stamford, CT 06906

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

DTNH22-09-D-00138, Task Order 7

12. Sponsoring Agency Name and Address

Office of 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; 05/15/13-011/14/15

14. Sponsoring Agency Code

15. Supplementary Notes
Amanda Kelley, PhD and Russell Pierce, PhD were the NHTSA Task Order Managers.

16. Abstract
This project evaluated the impact of the PC-based Risk Awareness and Perception Training (RAPT) program on young driver crashes and traffic violations. Young drivers 16 to 18 years of age were recruited immediately after they passed the on-road driving exam at six California Department of Motor Vehicles licensing offices. Participants were assigned to a group that completed the RAPT program or to a comparison group that received a pre-test but did not receive any training. A total of 5,251 young drivers participated in the project and had their crash and violation records tracked for 12 months post-licensure. An analysis of group equivalency revealed the group assignment paradigm was effective in producing equivalent groups. Analyses of pre-test and post-test data included in the RAPT program showed substantial improvements in trainee performance, which suggests participants attended to the training materials. Crash analyses did not show an overall main effect for treatment, but there was a significant treatment by sex interaction effect. Analyses were then conducted for males and females separately to explore this interaction. The results showed a significant treatment effect for males, $\chi^2(1, n = 2743) = 5.517$, $p =.019$, but not for females $\chi^2(1, n = 2447) = 0.553$ $p = .457$. RAPT-trained males showed an approximately 23.7% lower crash rate relative to the male comparison group. For females, the RAPT group had an estimated 10.7% higher crash rate than the comparison group, but this increase was not statistically significant. Researchers used Cox regression analysis to evaluate the number of weeks after licensure at which each studied driver had their first crash (time to first crash). None of the overall models were statistically significant. Thus, the hypothesis that RAPT had an effect on time to first crash could not be confirmed. None of the analyses of traffic violations demonstrated any association with the RAPT treatment. The results of this study provide perhaps the first encouraging evidence that brief, computer-based training interventions can have a positive influence on driving safety for newly licensed teen drivers. Further research is needed to clarify the uncertainties arising from this study, particularly related to the lack of effectiveness of RAPT on female crash rates, and to assess how best to employ hazard perception training using a program such as RAPT in the driver training process.

17. Key Words
Novice Driver
Field Test
Computer Training
Hazard Anticipation
Newly licensed Drivers
Driver education
Young Driver Crashes
Risk Perception

18. Distribution Statement

Document is available to the public from the National Technical Information Service www.ntis.gov.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	68	

Form DOT F 1700.7 (8-72)

ACKNOWLEDGEMENTS

The authors are grateful for the cooperation received from the California Department of Motor Vehicles (DMV) which made this study possible. Particular thanks go to Bayliss Camp, Michael Gebers, David DeYoung, and Robert Hagge of the DMV Research and Development Branch, and Graciela Pelayo-Hallum and Derek Oliver of the Field Operations Division for their inputs and tireless support. We also appreciated the excellent work in the field offices by the management and technicians who recruited participants and conducted the data collection. Finally, our thanks go out to the more than 5,000 newly licensed drivers and their parents/guardians who volunteered for the program.

Researchers adapted the versions of the Risk Awareness and Perception Training (RAPT) and comparison programs used for this study from the original RAPT produced by the University of Massachusetts at Amherst's Arbella Human Performance Laboratory directed by Donald Fisher. Left Brain Games, Inc., of Torrington, Connecticut, ably executed the programming for the RAPT revision and data collection modules.

While this study could not have been accomplished without the assistance of these organizations and individuals, the findings and conclusions presented herein are the responsibility of the authors and do not necessarily reflect the opinions or policies of the other project participants.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Executive Summary

Background and Objectives

An analysis of fatal crash data up to 2013 showed the fatal crash rate per mile driven is nearly twice as high for 16- and 17-year-olds as it is for 18- and 19-year-olds (Insurance Institute for Highway Safety, n.d.).This suggests the presence of one or more deficiencies during the first months of licensure that lead to an increased crash risk for newly licensed teen drivers. Analyses of police crash reports by McKnight and McKnight (2003) showed that the main failures leading to crashes among drivers 16 to 19 years old included inadequate visual scanning (ahead, to the sides and to the rear—43.6%), poor attention maintenance (23.0%), and inappropriate speed management (20.8%). Research has also shown that this group has poor anticipation of the existence and position of unexpected hazards (Pradhan et al., 2005) and does not control the speed, acceleration, and position of their vehicles well for hazard avoidance (Fisher et al., 2002; Sagberg & Bjørnskau, 2006).

One way to respond to these teen crash problems is to train new teen drivers to counter the deficiencies in their driving skills. One such training program that addresses hazard anticipation and poor scanning issues is the Risk Awareness and Perception Training (RAPT) program (e.g., Fisher et al., 2010). Researchers have conducted multiple studies of the impacts of various versions of the RAPT program approach using novice drivers in simulators (Pollatsek, Narayanaan, Pradhan, & Fisher, 2006) and in the field (Pradhan, Pollatsek, Knodler, & Fisher, 2009). These studies found that RAPT increased the likelihood that newly licensed drivers would anticipate hazards correctly as measured by the position of their gaze as they drove.

The current research effort had two phases. Phase 1 objectives included:

- Updating the packaging and delivery of the RAPT training so the resulting RAPT product could be confidently delivered to a wide audience.
- Training a large sample of newly licensed teen drivers using the updated RAPT program and collecting information from a similar group of newly licensed teens who did not receive the training.

Phase 2 objectives included:

- Constructing a database that contained participant demographics, RAPT training data, and one year of post-training driver record data (crashes and traffic violations).
- Analyzing the data to answer the following research questions:
    - During the first year of driving, did drivers who completed the RAPT program have fewer crashes or traffic violations on their driving records than the comparison group?
    - Is there a differential effect of RAPT training by sex as measured by number of crashes; types of crashes; and violations?
    - Does individual performance on the pre-test and post-test included in the RAPT program correlate with crashes and citations?

Method

Updating the program. In order to enhance the flow and appearance of the training and increase the stability of its operation, the entire training package was re-programmed and streamlined with higher quality graphics.

Site Selection. The researchers selected California as the site for this research because the California Department of Motor Vehicles (DMV) had an excellent and interested research department and an extensive number of licensing offices in which RAPT could be given to newly licensed young drivers. DMV selected six offices, three in the Los Angeles ("South") area and three in the San Francisco Bay ("North") area, with a high-flow of young drivers applying for first-time driver licenses.

Participants. Participants included 5,251 drivers 16 to 18 who had just passed their on-road driving exams for provisional or unrestricted license (first licenses) at one of the six selected California DMV field offices.

Training Procedure. A DMV technician acted as proctor and asked the participant to sit at a computer and read the first screen that explained the study. Once the participant consented and typed in his/her driver's license number, the RAPT or comparison (pre-test only—no training) program began. Proctors ran either the RAPT or comparison program on alternate weeks such that all participants in an office for a given calendar week completed either the RAPT or comparison program.

Data. Driver license number information and scores from measures generated within the RAPT and comparison programs were obtained from the desktop computers for all participants. License and RAPT score information from all offices were combined into a single file that was then merged with crash and violation data obtained from DMV's Driver Record Master (DRM) files for a period of 12 months post-licensure. Crash data were obtained without knowledge of participant treatment group assignment. Property damage only (PDO) and injury crashes contained in the DRM were obtained from the California Highway Patrol (CHP) and through self-reports to the State (i.e., by a crash-involved driver and/or an insurance company). There were no fatal crashes and too few injury crashes to analyze independently. The DRM provided only a total count of reportable violations and, therefore, researchers could not investigate individual violation types.

Analysis. The project involved several analyses including tests of group equivalency, analysis of RAPT pre-test/post-test scores, crash analyses (frequency and time to first crash), and analyses of traffic violations. Analyses focused on differences between the RAPT and comparison groups and by other covariates of interest.

Results

Group Equivalency. Logistic regression was used to evaluate group equivalency between the RAPT treatment and comparison conditions based on the available licensing and demographic information. The overall chi square statistic for the logistic model was not significant, χ^2 (15, N = 5190) = 22.031, p = .11. This suggests the paradigm used to assign drivers to the experimental and comparison groups, although quasi-random in nature, was sufficient to produce equivalent groups.

RAPT pre-test/post-test scores. Results showed minor differences in the pre-test scores for the RAPT and comparison groups, but the differences were very small and not practically meaningful. An analysis of the average number of primary hazards correctly identified across all nine scenarios for the RAPT group showed a large increase from the pre-test (M = 1.98, SD = 1.76) to post-test (M = 6.77, SD = 2.13), t(2540) = 110.24, p < 0.001, r = .38. This increase in performance is consistent with prior research (Pradhan, 2009) and likely suggests that participants attended to the RAPT training. Without post-test data for the comparison group, however, it is not known if scores increased simply because of repeated testing.

Crash Analyses. Table ES-1 shows the average number of crashes per driver for the first 12 months post-licensure by treatment group, sex, and age. Researchers analyzed the crash data using a Poisson regression approach. The results did not show a main effect for treatment, but did exhibit a significant treatment by sex interaction. Analyses were then conducted for males and females separately to explore this interaction. The results showed a significant treatment effect for males, χ^2(1, n = 2743) = 5.517, p =.019, but not for females χ^2(1, n = 2447) = 0.553 p = .457. RAPT-trained males showed an approximately 23.7% lower crash rate relative to the male comparison group, and this reduction was statistically reliable. For females, the RAPT group had an estimated 10.7% higher crash rate than the comparison group, but this increase was not statistically significant. The RAPT pre-test, post-test, and residualized change scores showed no relationship to crashes.

Table ES-1. Mean Crashes per Driver 12 Months Post-Licensure by Treatment and Sex

Treatment Group	Sex	Age (years)	Crash Rate	*SE*
Comparison	Male	16	0.080	0.0142
		17	0.129	0.0198
		18	0.116	0.0138
	Female	16	0.096	0.0168
		17	0.052	0.0141
		18	0.083	0.0132
RAPT	Male	16	0.096	0.0154
		17	0.074	0.0155
		18	0.075	0.0105
	Female	16	0.094	0.0145
		17	0.063	0.0136
		18	0.100	0.0145

Researchers used Cox regression analysis to evaluate the number of weeks after licensure at which studied drivers had their first crash (time to first crash). Multiple models were assessed for the best fit, but none of the overall models reached statistical significance ($ps > 0.05$). Thus, it could not be concluded that RAPT had an effect on time to first crash.

Violations. None of the analyses of traffic violations demonstrated any association with the RAPT treatment.

Discussion

The study achieved its process objectives, and the participant assignment process employed produced a dataset without meaningful bias between those who did and did not receive RAPT training. This dataset supported comparative analyses of the effects of RAPT on the crash and violation records of trained and similar untrained drivers during the initial year of their unsupervised driving careers.

The crash results indicated that RAPT training was associated with a statistically significant crash rate reduction of 23.7% for males. Interestingly, females exposed to RAPT showed a higher crash rate relative to females who did not complete the training, but this difference was not statistically significant. A differential effect on the sexes is not inconsistent with previous research focused on the teen driver. Although the precise causes of the differential response to RAPT by sex are currently unclear, factors dealing with both traffic safety and computer-based learning could have played a role. The analyses of time to first crash showed no effect of RAPT on time to first crash.

Overall, the results of this study provide perhaps the first encouraging evidence that brief, computer-based training interventions can have a positive influence on driving safety for newly licensed teen drivers even if only for males. Given the size of the sample included in this study, the single State venue, the fact that the RAPT implementation studied used only a limited number of selected scenarios, and the experimental limitations of the study, the reader must exercise caution when generalizing these findings. Nevertheless, the potential importance of the existence of some positive crash-based results for guiding future research and development with respect to the driver training process cannot be overlooked. Further research is needed to clarify the uncertainties arising from this study, particularly related to the lack of effectiveness, and even potential detrimental effects, of RAPT on female crash rates. Once a better understanding of the effects of RAPT is achieved, it would be useful to assess how best to employ hazard perception training using a program such as RAPT in the driver training process.

Evaluation of the Safety Benefits of the Risk Awareness and Perception Training Program for Novice Teen Drivers

Background

An analysis of fatal crash data up to 2013 showed the fatal crash rate per mile driven is nearly twice as high for 16- and 17-year-olds as it is for 18- and 19-year-olds (Insurance Institute for Highway Safety, n.d.). This suggests the presence of one or more deficiencies during the first months of licensure that lead to an increased crash risk for newly licensed teen drivers. Analyses of police crash reports by McKnight and McKnight (2003) showed that the main failures leading to crashes among drivers 16 to 19 included inadequate visual scanning (ahead, to the sides and to the rear—43.6%), poor attention maintenance (23.0%), and inappropriate speed management (20.8%). Research has also shown that this group has poor anticipation of the existence and position of unexpected hazards (Pradhan et al., 2005) and does not control the speed, acceleration, and position of their vehicle well for hazard avoidance (Fisher et al., 2002; Sagberg & Bjørnskau, 2006).

One way to respond to these teen crash problems is to train new teen drivers to counter the deficiencies in their driving behavior suggested by crash research. To this end, a variety of new driver training programs have been developed, in both the United States and abroad, including computer-based training (e.g., DriverZED produced by the AAA Foundation for Traffic Safety), as well as more advanced simulation programs such as those developed in Australia (e.g., DriveSmart). Such programs typically focus on advanced skills such as hazard anticipation and risk assessment, and have been evaluated for effectiveness in the simulator (Fisher et al., 2002; Regan, Triggs, & Godley, 2000) and on closed courses (Chapman, Underwood, & Roberts, 2002). The results of these studies often indicated the training was at least temporarily effective at improving scanning skills, but one study showed the effectiveness dissipated with time after training (Chapman, Underwood, & Roberts, 2002).

More recently, researchers at the University of Massachusetts at Amherst developed the Risk Awareness and Perception Training (RAPT) computerized program (Pollatsek, Narayanaan, Pradhan, & Fisher, 2006; Pradhan, Pollatsek, Knodler, & Fisher, 2009; Fisher, et al., 2010). The RAPT program evolved over time from a simple PowerPoint-based presentation of hazardous driving scenarios to more intricate sequences of actual photographs of potential problem situations from the perspective of a driver inside a vehicle. Through all of its versions, the RAPT program presented only one potentially hazardous situation at a time. In the original RAPT implementation, trainees saw a two-dimensional overview of a hazardous situation and had to "drag and drop" yellow ovals to areas on the screen where a potential threat might be located (e.g., in front of a truck) and red circles to locations on the plan view where they expected the threat to materialize (e.g., to the left and front of a truck). Trainees were then told the correct locations of the two markers.

The second generation of the RAPT program had trainees view sequences of photographs of hazardous situations. The sequences included views taken every few seconds from the driver's seat of a vehicle as it approached and passed a hazardous situation. The trainee had to use the mouse to move the cursor and click the locations in each photograph where he or she would be looking if driving. This version of the program had the correct scan areas defined within the program, and users received feedback that indicated if each hazard was correctly identified, although the exact hazard was never highlighted during the feedback. If the trainee gave an incorrect answer, the program repeated the training overview of the situation and sequence of photographs up to five times until he or she clicked on the correct target area in the photo sequence. If the individual did not succeed after five times through the training, the program moved on to the next training scenario without identifying the target hazard.

Researchers have conducted multiple studies of the impacts of the RAPT approach using novice drivers in a simulator (Pollatsek, Narayanaan, Pradhan, & Fisher, 2006) and small-scale studies on live roadways with researchers in the vehicle and the participants equipped with eye marker cameras. (Pradhan, Pollatsek, Knodler, & Fisher, 2009). These studies found that RAPT increased the likelihood that newly licensed drivers would anticipate hazards correctly as measured by the position and duration of their gaze as they drove. The research documented a positive effect of RAPT training both immediately after exposure to the training program and one week later. The effect was present both in a driving simulator and on the roadway, and was evident both in scenarios which were similar and dissimilar to trained situations. In one study of the program's effectiveness, the trained group fixated on the critical region of a simulated scene 77.4% of the time whereas a control group fixated on the critical region only 40.0% of the time (Fisher et al., 2007). During on-road data collection, the trained group fixated on the hazards correctly 64.4% of the time compared to 37.3% correct fixations for the untrained group.

Overall, researchers had examined the effectiveness the PC-based RAPT approach using various surrogate measures such as gaze position while research subjects drove a simulator or on road in controlled tests. The success of RAPT in these research environments led to an interest in determining if the RAPT approach could be effective in reducing crashes among newly licensed teen drivers. The current study exposed a large sample of newly licensed teen drivers to the RAPT program and tracked their subsequent crash histories for a one-year period in order to provide the most in-depth examination to date of the potential for the RAPT program to reduce teen driver crashes.

Objectives

The current research effort was designed as a two-phase project. The objectives during Phase 1 were to:

- Update the packaging and delivery of the RAPT training so the resulting RAPT product could be confidently delivered to a wide audience.
- Train a large sample of newly licensed teen drivers using the updated RAPT program and collect information from a similar group of newly licensed teens who did not receive the training.

The objectives during Phase 2 were to:

- Construct a database that contained participant demographics, RAPT training data, and one year of post-training driver record data (crashes and traffic violations).
- Analyze the data to answer the following research questions:
    - During the first year of driving, did drivers who completed the RAPT program have fewer crashes or traffic violations on their driving records than the comparison group?
    - Is there a differential effect of RAPT training by sex as measured by number of crashes; types of crashes; and violations?
    - Does individual performance on the pre-test and post-test included in the RAPT program correlate with crashes and citations?

Updating the RAPT Program

While the initial versions of the program proved effective at changing scanning behaviors in the lab and in small sample on-road studies, a large-scale field test required a more refined, robust, and stable version of the program that was compatible with the widest possible range of computers. To be consistent with the original RAPT training, the new program developed for the current study retained the same three modules—pre-test, training, post-test—the same response approach (participants were told to click where they would look), and the same photographs used in the prior version of RAPT. Also, the pre-test, training, and post-test modules each presented the same nine scenarios in the order that had been used in previous versions of the program. The program paused after each scenario, and the participant had to click "Next" to move on. Table 1 provides a brief description of each scenario and its primary hazard or target area—the area on which a participant would have to click to provide a correct response. The new RAPT retained the same primary hazard/target areas as in the earlier versions of RAPT.

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
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 (bus) obscures view of overhead traffic signal and oncoming traffic.	Additional traffic light not obscured by large vehicle
8	Two lane roadway approaching left turn; hill crest 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 block view of traffic approaching from left.	Top of hill/bend on left where oncoming traffic could emerge

In order to enhance the flow and appearance of the training, the entire training package was re-programmed in Adobe Air, which is easier to maintain and more universally applicable than the Director software in which the older version had been programmed. Also, the two-dimensional training diagrams used to introduce each training scenario were replaced with higher quality graphics. Other revisions included limiting the number of repetitions of each of the training sequences to two in order to control the amount of time the training could take. The prior version had repeated each scenario up to five times if the trainee did not click on the trained hazard. 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 provided an explanation of the reason the situation was hazardous. Also, whenever a trainee clicked on a hazard correctly, the new version of RAPT displayed a congratulatory message and the red oval to highlight the hazard as a further reminder of the correct behavior.

Pilot testing revealed the functional changes made to the training (i.e., feedback on hazards, limiting to two repetitions) did not adversely affect the performance of trainees on the pre-test and post-test and appeared to enhance the training while also greatly reducing the average amount of time spent using the program. Feedback was only provided in the training portion of RAPT. Similar to the older RAPT programs, the pre-test and post-test did not provide accuracy feedback to the users. Figure 1 and Figure 2 provide an example of the updated training screens for Scenario 7. Figure 3 shows two of the photographs from Scenario 7 with the target area indicated by a red box. The red box was not visible to the trainees but was provided to the development programmers to establish the area for determining correct performance and recording hits/misses in the data output.

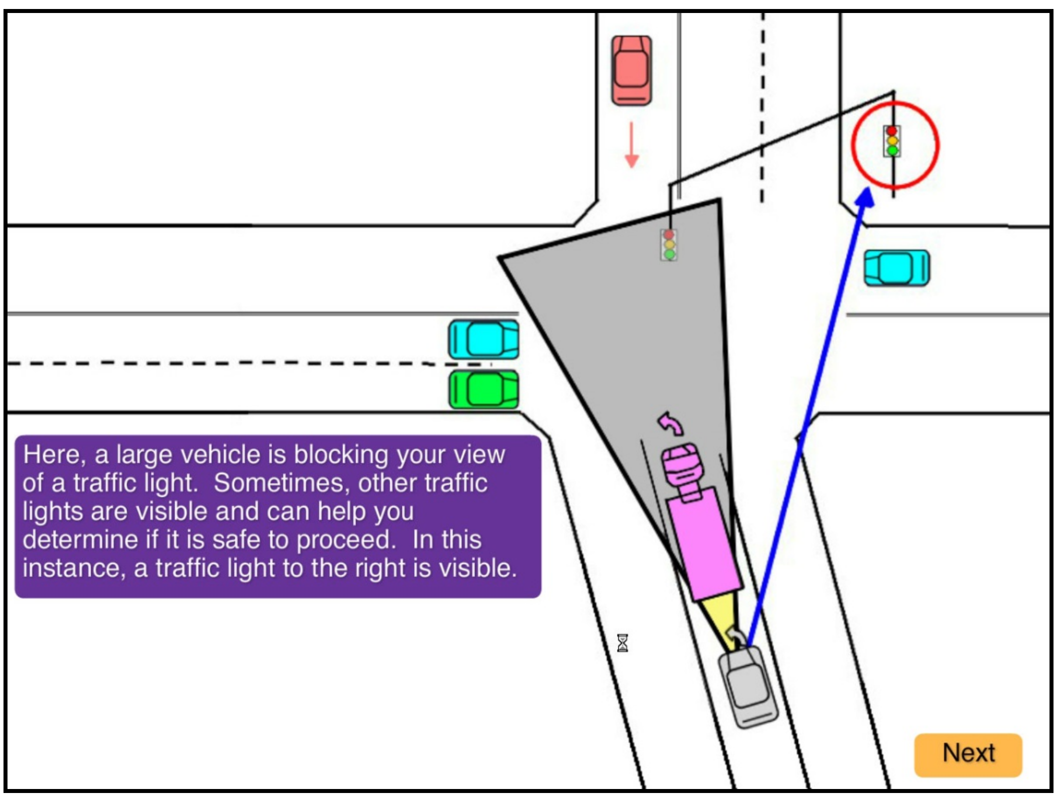


Figure 1. Example: Scenario 7 Training View—Screen 1

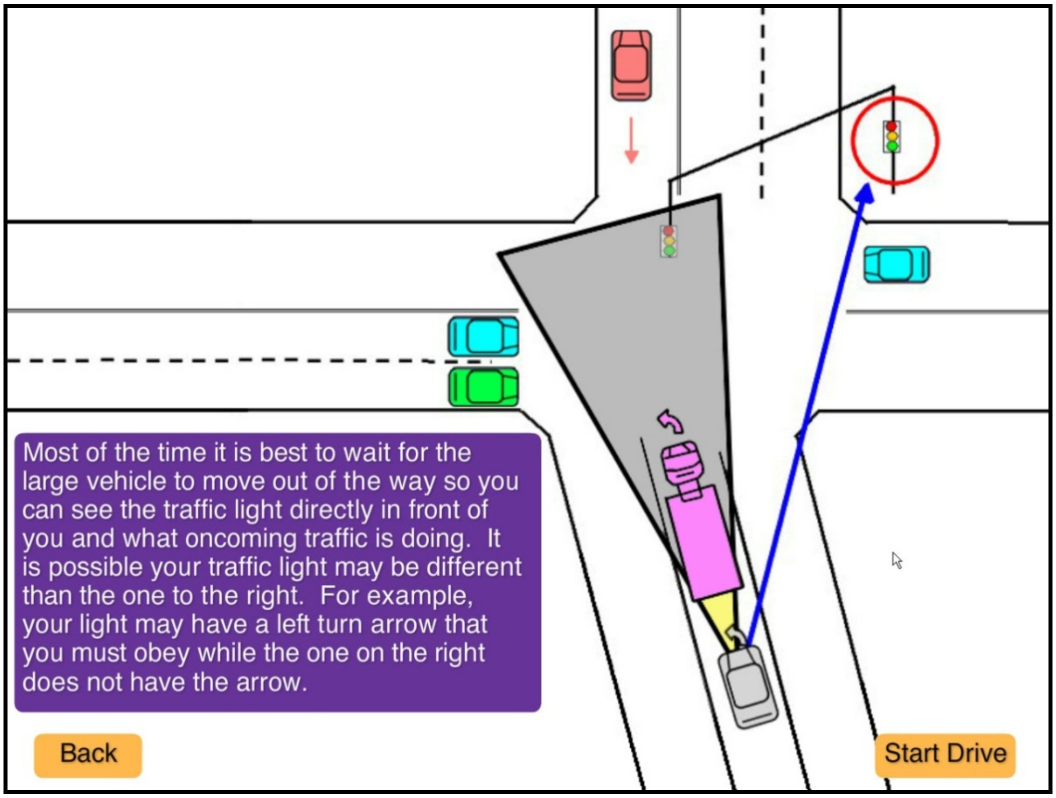


Figure 2. Example: Scenario 7 Training View—Screen 2

The other major changes to the program included the addition of a front-end "sign-on" module that captured the trainee's driver's license number and provided trainees with a description of the study that included a "check here if you agree to participate" consent function, and a data module that recorded training and performance data. In addition to the full RAPT program, a "comparison" version was created that contained only the front-end module that captured driver's license number, a description of the study, the consent check box, and the pre-test. This version without the training and post-test modules was designed for use by a comparison group and took 5 to 10 minutes to complete. No feedback on performance was provided to the comparison group. The pre-test was given to this group in order to examine the relationship of RAPT pre-test scores to crashes and violations and as a mechanism to assess group equivalency.

Given the results of small-scale studies of RAPT and similar programs using primarily behavioral measures, the next logical step was to conduct a field study to determine if this type of hazard anticipation training can actually reduce crashes and violations among newly licensed drivers. The remainder of this report covers the methods and results of a field study that sought to train a large sample of newly licensed drivers with the updated version of the RAPT program.

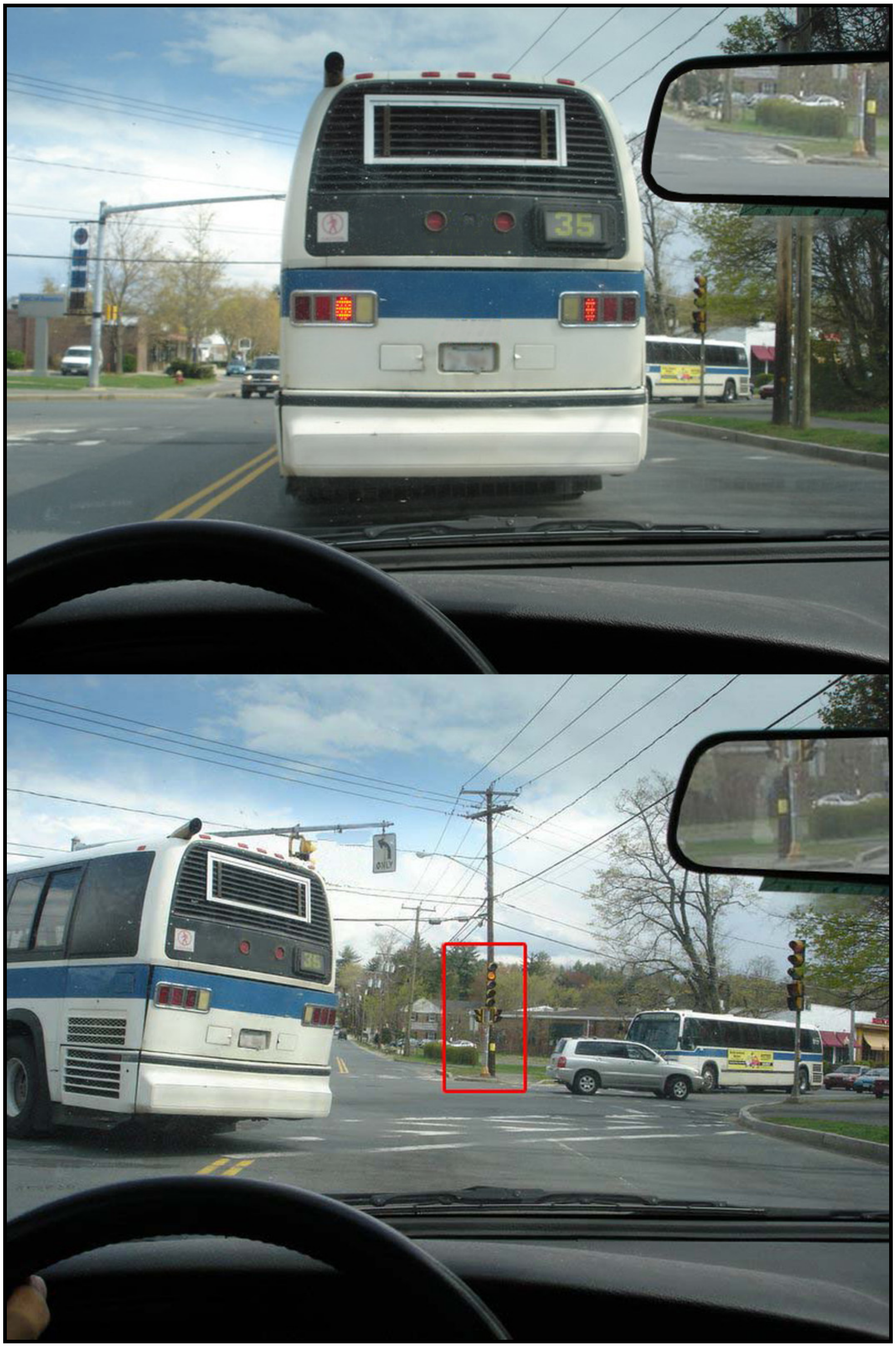


Figure 3. Example: Scenario 7 Photographs

Method

Site Selection

The researchers selected California as the site for this research because the California Department of Motor Vehicles (DMV) had an excellent and interested research department and an extensive number of licensing offices in which RAPT could potentially be given to newly licensed young drivers. California also had a "typical" and well-established graduated driver licensing (GDL) system (e.g., Masten & Hagge, 2004) including a requirement for a 6-month learner permit in the first stage (Level 1); independent driving with 12-month nighttime and passenger restrictions in the second stage (Level 2); and full unrestricted licensure in the third stage (Level 3). To apply for a learner permit, a teen was required to be 15½ years old and either have completed both driver education (30 hours of classroom-based learning) and driver training (6 hours of behind-the-wheel training), have completed driver education and enrolled in driver training, or simply be at least age 17½ years old. A teen could apply for an unrestricted license at age 18 without completing driver education or the GDL process. After researchers determined California was the site of choice, the NHTSA Regional office contacted the DMV to set up an initial meeting. At the first meeting, it was agreed to conduct a pilot study to determine the likely refusal rate of newly licensed drivers approached on-site at DMV field offices. Excessive refusals would undermine the validity of any data collection by creating a biased sample.

Feasibility Study

Researchers conducted the RAPT pilot study at two DMV field offices in the Sacramento area. The pilot test focused on determining the likely project participation rate of newly licensed drivers 18 or younger who had just passed the on-road drive test and received either provisional licenses or unrestricted licenses (if the drivers were 18 years old). The young driver and the person who brought the driver to the office (e.g., parent, guardian, sibling, driving instructor) participated in a semi-structured discussion concerning whether the driver would be willing to take the training if it were available and how peers might react if the training were offered at the DMV.

Overall, almost 95% of young drivers who participated in the discussions at the two DMV offices indicated they would take the training if it were available at the DMV. Some, however, expressed time concerns if the training took too long as many had left school to go to the DMV office for their license exam and drive test. The discussions also revealed that $30 would be sufficient to get most drivers to stay if the training took 30 minutes or less to complete. The great majority of the adults accompanying the teen drivers had no issue with staying after the drive test for the young drivers to complete the training, particularly since it was a safety-related program. The discussions also revealed that almost all people were willing to receive a check in the mail as a means of payment for participation (rather than cash on the spot) and were willing to provide their driver license numbers to permit the study to follow their driving records as long as no identifying information would ever be revealed as part of the study. Researchers also talked with the DMV technicians who participated in the pilot study and determined the developed protocol (described below) for identifying and referring drivers to the study would not hinder normal work activities.

DMV Field Office Selection

Given the positive results of the pilot study, DMV began the process of selecting appropriate field offices. The primary selection criteria were the rates of young drivers 16 to 18 applying for first-time driver licenses, the availability of sufficient space in the office to set up a study room, and proximity to other high-flow-rate offices to allow for easier project management. The higher flow rate offices generally conducted on-road licensing examinations for 10 to 50 young drivers a day depending on the time of year. DMV noted that a relatively large percentage, 50% or higher at times, failed the on-road licensing test, which meant the number of young drivers eligible for this project would normally be in the 5 to 25 range on any given day at a high-flow-rate office.

Based on the above rates in combination with the likely participation rate information from the pilot test, researchers estimated an office would likely yield 10 to 15 participants per day if the project started during the summer when higher numbers of young drivers applied for first-time licenses. Given the desire to have a sample of 5,000 young drivers (2,500 RAPT-trained and 2,500 controls), researchers estimated the project would need 333 to 500 office days. DMV therefore decided to select six offices, three in the San Francisco Bay (North) area and three in the Los Angeles (South) area, with a target of completing the field data collection in approximately 3 months.

Additional Meetings, Planning, and Training

Once DMV identified the offices, researchers held several more planning meetings with the DMV Field Operations Division and Research and Development Branch to finalize protocols for the project and prepare the field offices to use the new RAPT training equipment. Once DMV management approved all of the procedures, researchers visited each office to set up the dedicated RAPT training computers and train the field office staff who recruited the participants and served as the proctors for the training. Each DMV office assigned three to five staff to serve as proctors, with the staff members rotating from their normal duties to the proctoring duties on a daily basis. This approach kept the proctors engaged in both the study and their normal daily activities. When a DMV staff member served as a proctor, he or she did not conduct any other office work outside of the study room. The study rooms had to be fully staffed at all times to ensure no qualifying young drivers were missed. A subcontract agreement between Dunlap and Associates, Inc. and the DMV covered the participation of DMV office personnel in the project.

Training for the DMV technicians issuing the licenses focused on when to hand out recruitment flyers (immediately after an eligible driver passed the on-road examination), and where to direct potential participants who had questions (to the study rooms). DMV field office technicians had to check the birth date of each person who passed the on-road examination to determine if the person qualified. Researchers provided the technicians laminated cards that indicated the person qualified if he or she was “born on today’s date in 1992 or after” (i.e., was less than 19 years old at the time of data collection).

Proctor training included direction on how to operate the computers running the RAPT and comparison programs, when to use each of the programs, responses to frequently asked questions, and what to do if certain situations with the computers or participants arose. The proctors were also responsible for making sure participants entered their license numbers correctly and understood that the address they entered into the RAPT program's opening module was where they would receive the check in payment for participation.

Participants

Participants included drivers 16 to 18 who had just passed their driving test for provisional or unrestricted licenses (first licenses) at one of the six selected California DMV field offices. The DMV staffs at the six participating offices processed a total of 5,251 participants for whom study data were potentially available. A small number of these participants were lost for final analyses due to various issues discussed in the results section later in this report. Across all offices combined, the participation rate for newly licensed drivers approached was estimated to be almost 78%.

Study Duration

Data collection began in the South area DMV field offices on April 20, 2011. Data collection efforts started the following week in the North area offices. Members of the research staff oversaw the first two days of data collection in each office to ensure the proper execution of the study protocols by all proctors. Participant recruitment and training began as planned and continued in all six DMV offices until September 30, 2011. Driver record data were obtained for all participants covering the time from the date their driver licenses were issued up to October 1, 2012. This provided at least one year of post-license driving for all participants.

Materials

Computers. The RAPT program ran on Hewlett-Packard desktop computers using the Windows XP operating system. Each desktop computer included a 17-inch monitor, full keyboard, and mouse. The study provided four identical computers to each of the six DMV offices. The computers were not connected to the DMV network or the Internet.

Revised RAPT Program and Comparison Program. The only icons visible on the computer desktop were for the revised RAPT program and the comparison group testing program. A yellow triangle icon with the label "Yellow Triangle" represented the RAPT program, and a red square icon with the label "Red Square" represented the comparison testing program. The proctors opened each program to the consent screen before participants arrived so that only the proctors actually saw the desktop. The RAPT and comparison programs operated identically until the participant completed the pre-test. At that point, the comparison group's program ended while the RAPT training program continued to the training and post-test modules.

Recruiting Flyers. Each day, the proctors provided the DMV technicians at the customer service counter who processed driving test applicants with a supply of exactly 50 recruitment flyers. When a qualifying driver passed the on-road examination, the DMV technician handed the person a recruiting flyer that described the research study. The flyers were 8½ x 11 inches printed in black ink on bright orange copy paper that matched the color of signs throughout the office directing people to the study rooms. Figure 4 contains a black and white version of the recruiting form.

Congratulations!

Because you passed your driving test, you can earn $30 for volunteering less than 30 minutes of your time to participate in a safety research study.

- **See the person at the computer stations in the study room marked by an orange sign.**
- **You must be 18 or younger.**

The research project is sponsored by the National Highway Traffic Safety Administration (NHTSA).

No California DMV or State money is involved in this study.

Dunlap and Associates, Inc. 110 Lenox Avenue, Stamford, Connecticut 06906-2300 (203) 323-8464
FAX (203) 964-0799

Figure 4. Recruiting Flyer

Exit Flyer. In the event an individual requested more information about the study after completing participation, proctors provided an exit flyer that contained a very brief summary of the study and contact information for the researchers. The flyer did not contain any information about differences in the computer programs an individual may have completed. Appendix A contains a copy of the exit flyer.

Proctor Tally Forms. At the end of each day, proctors completed a tally form that included the number of participants who completed a program on each of the study computers. Proctors tallied and summed these numbers to provide a total count of participants completing the program for the day. Proctors then retrieved and counted the remaining recruitment flyers from the DMV counter technicians to determine how many flyers technicians handed out to eligible participants. Researchers used these counts to calculate an approximate participation rate for that particular day by dividing the number of participants completing the program by the number of flyers handed out by technicians. Appendix A contains an example of the proctor tally form.

Proctor Guidebook. Researchers provided each office with a proctor guidebook in a three-ring binder that contained a detailed outline of the procedures for the daily activities as well as electronic and hard copies of all recruiting and tally forms. The guidebook also included a calendar for each office that displayed a red square or yellow triangle for each date to notify proctors which program to run on that day. The proctor was responsible for ensuring that sufficient copies of all forms were available at all times.

Recruiting Procedure

Young drivers first had to pass the on-road examination to qualify for the study. After a driver passed the examination, the examiner directed the young driver to a pre-assigned DMV counter technician to complete the necessary paperwork for the new license. It was at this point that the technicians handed the young drivers the orange recruitment flyer. Potential participants read the flyers while they waited for the technician to complete their license paperwork, which included a temporary license (without a photo). In general, the flyer was sufficient to describe what the driver must do next for the study, but technicians often had to give short explanations such as, "the project is for a highway safety study, and you will receive $30.00 for participating." If they wished to participate, the newly licensed driver proceeded to the study room. If drivers or their parents/guardians had further questions, the technicians directed them to the study room where the proctor answered the questions.

Training Procedure

Proctors ran either the RAPT or comparison program for an entire week such that all participants in an office for a given calendar week completed the same program. The next week at the office, all participants completed the other program. The administration of the programs was counterbalanced such that the North area offices implemented one program on a given week while the South area offices implemented the other. Each office had a calendar indicating which program to administer each day. At the beginning of each day, proctors started the appropriate program on all four study-provided computers thereby allowing for the training of up to four participants simultaneously.

Once a participant entered the study room, proctors introduced themselves and checked the participant's temporary license to make sure it was issued on that day. If the license was not issued on that day, the person was not allowed to participate since it was not possible to know if the person had completed the training previously or if he or she was actually the individual named on the license. In addition, the proctor also re-checked the birth date on the license to ensure the person qualified for the study. The individual who drove the participant to the DMV was asked to wait outside the study room if possible in order to prevent any interference with the training.

After confirming the qualifications of a participant, the proctor asked the participant to sit at one of the four computers and to read the first screen that explained the study, why the license number was needed, and that the person would receive payment in the mail. The participant then had to use the computer mouse to click on a box to indicate agreement to participate. See Appendix A for a screenshot of this agreement to participate. The next screen asked for the person's license number and a mailing address for the payment. The proctor monitored the entry of these data to help ensure accuracy. Once the participant submitted this information, the RAPT or comparison program began. The programs operated without the need for any proctor intervention unless the individual had a question or the proctors noticed that the person was not following the directions. In general, the proctors did not interact with participants during their use of the program.

After the young driver completed the program, the proctor thanked him or her for participating and provided an exit flyer if the individual or the guardian requested more information.

Data Offload and Security

For security purposes, the computers in each office were not connected to any internal or external network. Therefore, research staff manually offloaded the check payment information and the program performance data files on a weekly basis. This required weekly visits to each office where the staff member offloaded the data to an external flash drive. The staff member then used a secure transmission channel to deliver the information to the home research office for processing. Researchers aggregated the data from the 24 computers to create analysis and check payment files.

Participant Payments

Researchers provided each weekly list of payee names and mailing addresses to a check fulfillment company that then mailed the $30 checks. Participants generally received their checks within two weeks of their participation date.

Data Sources, Processing, and Coding

Researchers obtained driver's license numbers and scores from measures generated within the RAPT and comparison programs (pre-test and post-test, as applicable) from the desktop computers for all participants. DMV used the driver's license numbers to access driver records from DMV's Driver Record Master (DRM) files for 12 months post-licensure. The driver record data was pulled without knowledge of participant training group assignment. RAPT training and score data from all offices were combined into a single file that was then merged with the crash and violation data.

Before attempting to access the crash and violation records, DMV personnel evaluated the accuracy of the driver's license number each participant had typed into the computer at the beginning of their session to ensure that it contained the correct number of digits, was not the number of an ID card, and that there was an actual record associated with the license number. Additional information (i.e., age at licensure, licensed in study period, licensed at study office) was also examined to ensure that the license number was associated with an individual who met the study inclusion criteria as a further check that the participant had not entered an incorrect license number. The data screen also removed data for any participant who had died during the first 12 months post-licensure.

The reference date (the date that the participant entered the study) was determined from the date of original driver's license issuance contained in the DRM. For the analyses presented in the following sections, the reference date was used to anchor a driver's subsequent history of crashes and violations in time.

Crash data contained in the DRM emanated from the CHP and through self-reports to the State (i.e., by a crash-involved driver and/or an insurance company). California Vehicle Code Section 16000 requires that all crashes resulting in an injury, fatality, or property damage above $750 be reported to the DMV. For the current study, crash data included property damage only (PDO) and injury crashes. Given the small sample size, there were no fatal crashes and too few injury crashes to analyze independently.

A count of all court-reported traffic violations was also included in the DRM. These violations included countable (e.g., speeding) and non-countable (e.g., broken tail light) violations, failure-to-appear in court for a traffic violation, and citations that were dismissed and/or masked after completion of traffic violators' school. In addition, both minor violations (e.g., failure to yield) and major violations (e.g., driving while intoxicated) were included. The date of violation, as opposed to the date of conviction, was used to reflect the date the event occurred. It is important to note that only a total count of reportable violations was available. As such, it was not possible to investigate individual violation types.

Analysis

For the results presented in the following sections, researchers analyzed data using IBM SPSS Statistics, Version 19. An initial set of simple descriptive analyses examined counts of participants by office and treatment condition and the response rate at each office. A second set of analyses focused on comparing pre-test scores for participants in the RAPT and comparison groups. Only the pre-test could be used for this analysis since the comparison group did not complete a post-test. Participants who did not complete all nine pre-test scenarios were excluded from these analyses. Researchers first used chi square tests to compare the numbers of participants in each treatment condition that correctly identified ("hit") or missed targets during each scenario on the pre-test. A chi square test was also used to examine the distribution of total number of targets hit during the pre-test by treatment group. An independent samples *t*-test was used to compare the mean number of targets hit during the pre-test for each treatment condition. Researchers used ANOVA to compare pre-test scores for RAPT training group dropouts to those who completed the entire RAPT program and the comparison group.

The next set of analyses focused on the RAPT training group's change in targets hit from pre-test to post-test. Only participants who had complete pre-test and post-test data were included in these analyses. Chi-square tests were used to examine the numbers of participants who hit/missed the target for each scenario on the pre-test and post-test. A t-test then compared the mean total targets hit (0-9) for the pre-test and post-test. Researchers used a repeated measure ANOVA to explore any differences in performance by sex and age for the pre- and post-tests. The results also include a brief set of descriptives for the amount of time it took participants to complete the various sections of the RAPT and comparison programs. Univariate ANOVA was used to examine differences in training times by sex and age.

Further analyses focused on demographic data obtained from the DRM for the participant groups in each treatment condition. All participants with valid DRM data, even those who did not complete the RAPT program, were included in these analyses. Descriptive results were calculated for participants in the RAPT and comparison conditions based on several demographic variables, including sex, age, field office at which they took their driving test, and day of the week to ensure that there were no significant differences between the RAPT and comparison groups on these variables. ZIP Code indices for mean crash and conviction rates were also calculated. These indices provide an estimate of average crash and conviction rates per postal ZIP Code over a 36-month period.

Logistic regression was used to evaluate group equivalency between the RAPT treatment and comparison conditions based on the available licensing, biographical, and demographic information. As a final assessment of group equivalency, receiver operating characteristic (ROC) curves were calculated for several logistic regression models containing various subsets of variables.

The primary study analyses focused on crash rates of groups based on treatment (RAPT-trained or comparison), sex, and age at licensure. Researchers analyzed the crash involvement data with a Poisson regression analysis comparing participants who received the RAPT training and the comparison group on the number of crashes during the first 12 months post-licensure. Researchers adopted an intent-to-treat approach that included all participants with valid crash data even if they dropped out before completing the assigned training/testing program. Unless otherwise specified, researchers assessed statistical significance by using two statistical tests: (1) the likelihood ratio chi square for testing both overall model fit and for differences between models with different subsets of variables, and (2) the Wald chi-square test for testing the regression parameters associated with individual effects within models.

To further evaluate the potential effect of RAPT training on subsequent driving records, researchers conducted analyses to evaluate whether a relationship existed between performance on the RAPT tests (number of scenarios in which the primary target was correctly identified) and crash rates for participants in the RAPT treatment condition only. Analyses included simple correlations of RAPT scores with crashes and violations as well as Poisson regression.

A second approach employed to evaluate crash risk between the RAPT treatment and comparison groups involved evaluating the number of weeks to first crash. Cox regression analysis provided inferential analysis of the weeks to first crash with the covariates of interest and corresponding interactions entered into the equation.

The study included a research question focused on whether RAPT had any association with driver traffic violations in the first year of licensure. Researchers conducted binary logistic regression analyses to examine violation rate differences based on treatment (RAPT-trained or comparison), sex, and age at licensure. Researchers also conducted a Cox regression analysis to determine if there were any differences in weeks to first violation for the covariates of interest.

Results

Participant Counts and Participation Rates

The total number of participants entering the study varied by office with the South area offices tending to have larger numbers of participants (61% of the sample) than the North area offices. The total in each office, however, was approximately evenly split between the RAPT and comparison groups. Across all offices, 50.7% of all participants were assigned to complete the RAPT program, and 49.3% were given the comparison program (Table 2).

Table 2. Count of Participants by Office and Training Group

Office		RAPT-Trained*	Comparison Group*	Total
South 1	Count	615	548	1163
	Office %	52.9%	47.1%	100.0%
	Total N %	11.7%	10.5%	22.2%
South 2	Count	707	757	1464
	Office %	48.3%	51.7%	100.0%
	Total N %	13.5%	14.4%	27.9%
South 3	Count	298	274	572
	Office %	52.1%	47.9%	100.0%
	Total N %	5.7%	5.2%	10.9%
South Subtotal	Count	1620	1579	3199
	Office %	50.6%	49.4%	100.0%
	Total N %	30.9%	30.1%	61.0%
North 1	Count	514	467	981
	Office %	52.4%	47.6%	100.0%
	Total N %	9.8%	8.9%	18.7%
North 2	Count	190	208	398
	Office %	47.7%	52.3%	100.0%
	Total N %	3.6%	4.0%	7.6%
North 3	Count	334	330	664
	Office %	50.3%	49.7%	100.0%
	Total N %	6.4%	6.3%	12.7%
North Subtotal	Count	1038	1005	2043
	Office %	50.8%	49.2%	100.0%
	Total N %	19.8%	19.2%	39.0%
Total	Count	2658	2584	5242
	Office %	50.7%	49.3%	100.0%

*Counts based on participants with at least partial computer-recorded data indicating that they at least began their assigned condition in the DMV office.

Researchers examined the participation rate in each office based on the total number of flyers distributed by DMV technicians versus the number of participants tallied by proctors as completing the program. The number of days tallied varies by office due to the staggered study start and the fact that some offices did not collect data on some days for a variety of reasons unrelated to the study (e.g., all driver license examiners were at training). Total counts of participants from the tally forms differ from the counts obtained from the training computer programs because proctors may not have tallied an individual if the person started the program but left before completing any portion of the program, if a recruiting form was lost, or if the proctor simply miscounted at the end of the day. Across all offices combined, the participation rate was 77.9%. The data showed notable differences in the participation rates by office with the highest participation rate achieved being 97.1% and the lowest 49.9% (Table 3). When combined by region, the South area offices showed a higher participation rate (87.7%) than did the North area offices (66.5%). Anecdotal data from the proctors and researcher observations suggested that most refusals resulted from time pressure (e.g., the prospective participant had to get back to school; the accompanying adult had to get back to work).

Table 3. Participation Rate by Office

Office	Participants*	Refusals	Days Tallied	Participation Rate
South 1	1151	34	114	97.1%
South 2	1428	186	114	88.5%
South 3	563	221	115	71.8%
South Subtotal	3142	441	343	87.7%
North 1	996	391	108	71.8%
North 2	402	403	101	49.9%
North 3	628	228	106	73.4%
North Subtotal	2026	1022	315	66.5%
Totals	5168	1463	658	77.9%

*Counts based on proctor tallies of participants completing the program.

Pre-Test Performance for RAPT-Trained Group Versus Comparison Group

Each scenario had only a single area on one of its photographs defined as the hazard area of interest. If the participant clicked on this area while the appropriate photograph was displayed, the computer program logged a “hit” in the database. For the total participant sample, pre-test hit rates ranged from a low of 5.9% for Scenario 4 to a high of 49.2% for Scenario 6 (Table 4). In general, the group that would go on to receive RAPT training performed slightly better than the comparison group on the pre-test, identifying the hazard of interest at a minimally higher rate in eight of the nine scenarios (Table 4). Only the differences for Scenario 6, $\chi^2(1, N = 5114) = 5.75$, $p = 0.016$, $\varphi_c = .03$, and Scenario 8, $\chi^2(1, N = 5114) = 4.46$, $p = 0.035$, $\varphi_c = .03$, were statistically reliable. However, the differences of 3.3 percentage points for Scenario 6 and 2.2 percentage points for Scenario 8 were not large and are unlikely to be of practical significance. Up to the point that the participants completed the pre-test, both groups received equivalent treatment.

Table 4. Percent With Primary Hazard "Hit" During Pre-Test Scenarios

Scenario	Comparison (n = 2,573)	RAPT-Trained (n = 2,541)	φ_c	Total (N = 5,114)
1	24.1	25.2		24.6
2	13.8	14.4		14.1
3	17.8	19.6		18.7
4	5.4	6.3		5.9
5	8.8	9.3		9.1
6	47.6	50.9*	.03	49.2
7	28.2	27.3		27.7
8	14.7	16.9*	.03	15.8
9	26.5	28.4		27.4

*RAPT-trained significantly higher than comparison, chi square $p < 0.05$.

Table 5 displays the number and percentage of participants in each training group by overall performance on the pre-test as measured by the sum of the hazards correctly identified across the nine scenarios. The table shows the two training groups had very similar percentages of participants at each performance level, but the slight differences did reach statistical significance, $\chi^2(7, N = 5114) = 14.42$, $p = .044$, $\varphi_c = .05$. The relatively small effect size, pattern of results, and equivalent treatment given the two groups by the proctors and computer programs suggest the differences were not operationally meaningful. It is important to note that across both groups only 19.1% of the participants correctly identified four or more of the nine hazards on the pre-test.

Table 5. Total Number of Pre-Test Hazards Hit by Training Group

		0	1	2	3	4	5	6	7+
RAPT Trained	Count	661	496	515	344	274	156	68	27
	Row %	26.0%	19.5%	20.3%	13.5%	10.8%	6.1%	2.7%	1.1%
Comparison Group	Count	730	498	496	398	247	120	58	26
	Row %	28.4%	19.4%	19.3%	15.5%	9.6%	4.7%	2.3%	1.0%
Total	Count	1391	994	1011	742	521	276	126	53
	Row %	27.2%	19.4%	19.8%	14.5%	10.2%	5.4%	2.5%	1.0%

An analysis of the average number of hazards hit across all nine scenarios (i.e., sum of hits across the nine scenarios) showed the participants in the RAPT-trained group (M = 1.98, SD = 1.76) had a slightly higher number of hits on average than did the comparison group (M = 1.87, SD = 1.70), $t(5112) = 2.38$, $p = 0.017$, $r = .03$. Exploratory analyses showed that this difference in means was at least partially attributable to the exclusion from the analyses of pre-test scores for people who did not complete the entire RAPT training program. A total of 122 participants assigned to RAPT training failed to complete the entire training/testing program, and 15 of the comparison participants failed to complete the pre-test. Of the 122 RAPT dropouts, 97 completed the pre-test before dropping out. Researchers analyzed the pre-test scores of these drop-outs relative to those who completed the entire training program and the comparison group that received no training. A total pre-test score could not be calculated for the comparison group dropouts since, by definition, none of these dropouts actually completed the entire pre-test, which was their only task.

As shown in Table 6, the RAPT dropouts had a notably lower mean score on the pre-test compared to the participants who completed the entire RAPT program and the comparison group, $F(2, 5208) = 12.84$, $p < 0.001$, $\eta^2 = 0.005$.

Table 6. Mean Pre-Test Hazards Hit by Training Group and RAPT Dropouts

	N	*Mean*	*SD*	*SE*
RAPT Completed	2541	1.98	1.76	.0348
RAPT Dropout	97	1.13	1.55	.1569
Comparison	2573	1.87	1.70	.0336

Combining the pre-test scores for the RAPT training dropouts with the completed RAPT training group's pre-test scores reduced the overall mean to 1.95 ($SD = 1.76$). Analysis showed the mean for this intent to treat group (the combination of those assigned to RAPT who dropped out and those who completed RAPT) was not significantly different than the comparison group's mean, $t(5209) = 1.75$, $p = 0.08$. The information collected during the study does not provide an explanation for the significantly lower pre-test scores of the dropout group or the reasons why they chose not to complete the program.

RAPT Training Group Change in Performance from Pre-Test to Post-Test

Researchers examined the counts of hazard hits in each scenario for the RAPT group for the pre-test and post-test to determine the impact of the training on identification of the specific hazards covered by the training. The analysis showed substantial, and significant ($p < 0.001$), increases in hazard identification from the pre-test to the post-test for all nine of the scenarios (Table 7). An analysis of the average number of hazards hit across all nine scenarios (i.e., sum of hits across the nine scenarios) showed a large increase in correct identification of the defined hazards from the pre-test ($M = 1.98$, $SD = 1.76$) to post-test ($M = 6.77$, $SD = 2.13$), paired samples $t(2540) = 110.24$, $p < 0.001$, $r = .38$.

Table 7. Percent with Defined Hazard "Hit" for Pre-Test and Post-Test Scenarios

Scenario	Pre-Test (*n* = 2,541)	Post-Test (*n* = 2,541)	*V*
1	25.2	68.5*	.59
2	14.4	70.0*	.71
3	19.6	57.9*	.54
4	6.3	82.0*	.87
5	9.3	77.2*	.82
6	50.9	88.7*	.58
7	27.3	94.2*	.81
8	16.9	64.0*	.63
9	28.4	73.9*	.63

*Post-test significantly higher than pre-test, McNemar chi square $p < 0.001$.

The change in scores was further examined by the sex and age of the participant. A repeated measures ANOVA examined pre-test and post-test scores by age and sex with pre/post as a within subjects variable. The mean number of targets hit for each age and sex group for the pre-test and post-test are provided in Table 8. The N for this analysis is slightly smaller than the previous analyses since sex and age information could not be obtained for all participants. The results showed significant between subjects main effects for sex, $F(1, 2504) = 104.50$, $p < .001$, partial $\eta^2 = .04$ and age, $F(2, 2504) = 24.386$, $p < .001$, partial $\eta^2 = .02$, but the age by sex interaction was not significant ($p > .05$). The within subjects results showed a large effect for the pre/post variable, $F(1, 2504) = 11{,}254.21$, $p < .001$, partial $\eta^2 = .818$. The two-way interactions of pre/post by sex and pre/post by age were not significant (*p*s > .05). The three-way interaction of pre/post by sex by age was statistically significant, $F(2, 2504) = 3.05$, $p = .047$, partial $\eta^2 = .002$, but the effect size was very small. These results indicate that males tended to perform better than females overall, and younger participants scored higher than older participants on average. The three-way interaction of sex, age, and pre/post scores demonstrates that there were some slight differences in how much scores changed for the various age and sex groups from the pre-test to post-test, but these differences are not particularly meaningful given the extremely small effect size.

Table 8. Mean Targets Hit for Pre-Test and Post-Test by Sex and Age

		Pretest				Posttest			
		Mean	*SD*	*SE*	*N**	*Mean*	*SD*	*SE*	*N**
Male	16	2.59	1.88	.10	371	7.32	1.82	.09	371
	17	2.53	1.90	.11	298	7.11	2.00	.12	298
	18	1.99	1.83	.07	635	6.93	2.03	.08	635
	Total	2.28	1.88	.05	1304	7.08	1.97	.05	1304
Female	16	1.95	1.63	.08	408	6.66	2.02	.10	408
	17	1.66	1.47	.08	302	6.61	2.21	.13	302
	18	1.43	1.50	.07	496	6.19	2.37	.11	496
	Total	1.66	1.55	.04	1206	6.45	2.22	.06	1206

*Only includes participants with sex, age, and complete RAPT scores available.

Computer Testing/Training Times

The amount of time spent on the pre-test, training, post-test, and total elapsed time for the training was of interest as a descriptor of the training process. It should be noted the total elapsed time was calculated from the moment the person first clicked the box to start the program to the point where the final "Thank You" screen appeared upon completion of the post-test. Therefore, the total elapsed time exceeds the sum of the pre-test, training, and post-test times because it includes the time to read the instructions and may include any pauses or breaks taken by participants. Training and post-test times were not applicable to the comparison group.

Table 9 shows the means and standard deviations of the pre-test, training, and post-test program sections by treatment group in minutes:seconds (mm:ss) format. Pre-test times were virtually the same for the two groups at around 3:30, which is not surprising since the only variance in a participant's time to complete the section would be due to how long it took the person to read the task description and click "Next" to continue. The computer paced the remainder of the pre-test. The post-test mean time for the RAPT group was 03:24 (*SD* = 00:14), which is similar to the pre-test time as expected since the two tests were identical. Actual time spent on training for the RAPT group showed a mean of 08:10 (*SD* = 02:01) with a minimum of 04:15 and a maximum of 25:04. The mean total elapsed time for the RAPT group was 17:32 (*SD* = 02:46) with a minimum of 10:44 and a maximum of 44:11. The mean total time for the comparison group was 05:29 (*SD* = 00:55) with a minimum of 02:56 and a maximum of 09:16.

Table 9. Mean Computer Program Section Times by Group

Treatment Group	Pre-Test		Training		Post-Test		Total Elapsed Time	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
RAPT	03:32	00:18	08:10	02:01	03:24	00:14	17:32	02:46
Comparison	03:31	00:19	N/A	N/A	N/A	N/A	05:29	00:55

Researchers then examined time spent on the training section of RAPT as a function of participant age and sex using univariate ANOVA. The results showed significant main effects for sex, $F(1, 2504) = 80.21$, $p < .001$, partial $\eta^2 = .031$, and age $F(2, 2504) = 4.39$, $p = .003$, partial $\eta^2 = .003$, but not their interaction ($p > .05$). A review of the mean training times in Table 10 reveals that males took about 43 seconds less on average to complete the training than did females. Younger males and females tended to take less time overall than their older counterparts. An examination of pre-test and post-test times revealed that times only varied by about one second for the sex and age groups, which was expected given the fixed duration nature of the tests. The total elapsed time only varied by the difference in training times mentioned above.

Table 10. Mean Training Section Times by Sex and Age

		Mean	*SD*	*SE*	*N*
Male	16	07:42	01:50	00:05	371
	17	07:48	01:48	00:06	298
	18	07:57	02:01	00:04	635
	Total	07:50	01:55	00:03	1304
Female	16	08:23	01:51	00:05	408
	17	08:34	01:57	00:06	302
	18	08:42	02:14	00:06	496
	Total	08:33	02:03	00:03	1206

Treatment Group Equivalency

Descriptive data were calculated for participants in the RAPT and comparison conditions based on several licensing and demographic variables, including sex, age, field office at which they took their driving test, and day of the week (see Table 11), to ensure that no meaningful differences existed between the RAPT and comparison groups on these variables. Researchers also calculated ZIP Code indices for mean crash and conviction rates. Table 11 presents the data separately for the intent to treat group (everyone assigned to RAPT for whom data were available) and only the set of participants who completed the study (excluding those who dropped out).

Researchers used logistic regression to evaluate group equivalency between the RAPT treatment and comparison conditions based on the available licensing and demographic information. A likelihood-ratio test evaluated the model fit with all variables entered simultaneously relative to an intercept-only model. Analyses used sex, field office where the driver's license was obtained, day of the week that the license was obtained, age at date of licensure, RAPT pre-test score, completion of pre-test, and average total convictions and crashes by ZIP Code as predictor variables (Table 12). The overall chi square statistic for the logistic model was not significant, χ^2 (15, N = 5190) = 22.031, p = .11. This suggests the paradigm used to assign drivers to the experimental and comparison groups, although quasi-random in nature, was sufficient to produce equivalent groups. ROC curves (Figure B1 in Appendix B) were calculated for several logistic regression models containing subsets of variables. The areas under the curve (AUC) for these models ranged from .524 through .537 (Table B1), further supporting equivalency between the RAPT and comparison groups. Perfect equivalence would have yielded a value of 0.5 and complete non-equivalence a value of 1.0.

Table 11. Demographic Characteristics of the RAPT and Comparison Groups

		Intent to treat group		Completed only (dropouts excluded)	
		RAPT	Comparison	RAPT	Comparison
Sex	Males	1,384	1,397	1,304	1,376
		52.10%	54.10%	52.00%	54.00%
	Females	1,274	1,187	1,206	1,171
		47.90%	45.90%	48.00%	46.00%
Age	16 years old	813	778	779	776
		30.90%	30.40%	31.00%	30.50%
	17 years old	631	620	600	615
		24.00%	24.20%	23.90%	24.10%
	18 years old	1,184	1,164	1,131	1,156
		45.05%	45.40%	45.10%	45.40%
Field Office	South 1	616	549	610	543
		23.20%	21.10%	24.30%	21.30%
	South 2	708	757	650	739
		26.60%	29.30%	25.90%	29.00%
	South 3	298	274	283	272
		11.20%	10.60%	11.30%	10.70%
	North 1	514	468	471	460
		19.30%	18.10%	18.80%	18.10%
	North 2	190	208	187	205
		7.10%	8.00%	7.40%	8.00%
	North 3	334	330	309	328
		12.60%	12.80%	12.30%	12.90%
Day of the Week	Monday	471	473	448	468
		17.90%	18.50%	17.80%	18.40%
	Tuesday	501	513	482	509
		19.10%	20.00%	19.20%	20.00%
	Wednesday	457	447	436	446
		17.40%	17.40%	17.40%	17.50%
	Thursday	611	586	580	584
		23.20%	22.90%	23.10%	22.90%
	Friday	588	543	564	540
		22.40%	21.20%	22.50%	21.20%
ZIP Code Indices per 100 Drivers (36 months)	Crashes	12.94	12.98	12.94	12.98
	Convictions	55.65	55.64	55.65	55.64
	Percent of total crashes involving injuries[1]	21.67%	22.32%	21.10%	20.80%

[1]No significant differences were observed between groups for percentage of injury crashes.

Table 12. Logistic Regression Results Testing for Group Equivalency

							95% CI for *OR*	
Predictors	*β*	*SE*	χ^2	*df*	*p*	*OR*	*LL*	*UL*
Constant	0.218	0.581	0.141	1	0.708	1.244		
Day of the week (ref: Friday)			1.565	4	0.815			
Monday	-0.083	0.089	0.885	1	0.347	0.920	0.774	1.094
Tuesday	-0.092	0.087	1.121	1	0.290	0.912	0.769	1.081
Wednesday	-0.050	0.090	0.307	1	0.579	0.952	0.798	1.134
Thursday	-0.026	0.083	0.100	1	0.752	0.974	0.827	1.147
Field office city (ref: North 3)			11.536	5	0.042			
South 1	0.161	0.106	2.309	1	0.129	1.175	0.954	1.447
South 2	-0.073	0.103	0.507	1	0.477	0.930	0.760	1.137
South 3	0.091	0.117	0.608	1	0.435	1.095	0.871	1.376
North 1	0.098	0.103	0.913	1	0.339	1.103	0.902	1.350
North 2	-0.108	0.131	0.670	1	0.413	0.898	0.694	1.162
Sex (ref: female)	-0.105	0.057	3.409	1	0.065	0.900	0.806	1.006
Pre-test completion (ref: incomplete)	0.531	0.332	2.549	1	0.110	1.700	0.886	3.263
RAPT pre-test score	0.032	0.017	3.799	1	0.051	1.033	1.000	1.067
Age	-0.011	0.034	0.106	1	0.744	0.989	0.925	1.057
Average crashes by ZIP Code	-0.045	0.030	2.230	1	0.135	0.956	0.902	1.014
Average conviction by ZIP Code	0.035	0.031	1.267	1	0.260	1.035	0.975	1.100

Log likelihood: χ^2 (15, N = 5,190) = 22.031, p = .11

Crash Rates 12 Months Post-Licensure

Table 13 presents the mean number of crashes during the first 12 months post-licensure for the RAPT and comparison groups. A separate examination of crash rates by sex and age for the RAPT and comparison groups is shown in Table 14. Due to the small number of crashes, researchers did not perform analyses of crash severity. The interested reader may review the frequencies of crashes by severity in Table B2 of Appendix B.

Table 13. Mean Crashes per Driver During the First 12 Months Post-Licensure

Treatment Group	Crash Rate	*SE*
Comparison	0.089	0.0066
RAPT	0.082	0.0061

Table 14. Mean Crashes per Driver by Treatment, Sex, and Age

Treatment Group	Sex	Age (years)	Crash Rate	*SE*
Comparison	Male	16	0.080	0.0142
		17	0.129	0.0198
		18	0.116	0.0138
	Female	16	0.096	0.0168
		17	0.052	0.0141
		18	0.083	0.0132
RAPT	Male	16	0.096	0.0154
		17	0.074	0.0155
		18	0.075	0.0105
	Female	16	0.094	0.0145
		17	0.063	0.0136
		18	0.100	0.0145

To determine if any statistically significant differences existed between the RAPT and comparison groups, researchers analyzed the crash data using a Poisson regression analysis that compared the two groups on the number of crashes during the first 12 months post-licensure. Prior to running the Poisson analysis, five models including treatment group, sex, age, and interactions were developed (see Table 15) to determine the best fitting model for the total number of crashes. These models were derived from predictions regarding the effectiveness of RAPT and potential interactions with sex and/or age. Other covariates including DMV office and participant ZIP Code were included in preliminary analyses but did not affect the significance of the models and were not retained. Final analyses focused on the most parsimonious models.

Researchers evaluated the statistical significance of the models using the likelihood ratio chi square test. The most simplistic model, Model E, only includes the effect of treatment. Table 15 shows the *p*-value for the intercept and treatment in model E and demonstrates that the overall model that only included treatment was not significant. Models A and B were both statistically significant, and Models A, B, and C displayed an obvious tendency for a significant interaction involving treatment by sex. Researchers therefore focused on Model B since it was statistically significant overall, parsimonious, and contained a significant treatment by sex interaction (see Table 16 and Table 17).

Table 15. Poisson Models for Number of Crashes[1]

	Poisson Regression Models				
	A	B	C	D	E
Intercept	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$
Treatment	$p = 0.437$	$p = 0.304$	$p = 0.310$	$p = 0.204$	$p = 0.196$
Sex	$p = 0.111$	$p = 0.088$	$p = 0.255$	$p = 0.224$	
Age	$p = 0.207$	$p = 0.239$	$p = 0.449$	$p = 0.462$	
Treatment X Sex	$p = 0.058$	$p = 0.047$	$p = 0 .032$		
Treatment X Age	$p = 0.526$	$p = 0.452$			
Sex X Age	$p = 0.057$	$p = 0.056$			
Treatment X Sex X Age	$p = 0.108$				
	$\chi^2 = 22.037$ $df = 11$ $p = 0.024$	$\chi^2 = 17.585$ $df = 9$ $p = 0.040$	$\chi^2 = 9.397$ $df = 5$ $p = 0.094$	$\chi^2 = 4.771$ $df = 4$ $p = 0.312$	$\chi^2 = 1.672$ $df = 1$ $p = 0.196$

Table 16. Model B Poisson Regression Test of Model Effects

Source	Type III Likelihood Ratio χ^2	*df*	*p*
Intercept	6,924.829	1	0.000
Treatment	1.056	1	0.304
Sex	2.906	1	0.088
Age	2.867	2	0.239
Treatment X Sex	3.956	1	0.047
Treatment X Age	1.586	2	0.452
Sex X Age	5.768	2	0.056

[1] Poisson analyses for all models were conducted with and without covariates (day of the week, field office, pre-test completion, pre-test score, average crash by ZIP Code and average violation by ZIP Code), and no differences were observed. To make the models more parsimonious, covariates were not included in the final models.

Table 17. Model B Poisson Regression Model Parameter Estimates for Total Crashes

Parameter	*B*	*SE*	Hypothesis Test Wald χ^2	*df*	*p*	Exp (*B*)	95% CI Exp(*B*) *LL*	*UL*
Intercept	-2.365	0.133	318.550	1	0.000	0.094	0.072	0.122
Sex (ref: female)	-0.159	0.169	0.888	1	0.346	0.853	0.612	1.188
Age (ref: 18 years)								
16 years	0.119	0.181	0.430	1	0.512	0.937	0.620	1.415
17 years	-0.512	0.231	4.928	1	0.026	0.625	0.373	1.047
Treatment (ref: RAPT)	-0.048	0.172	0.078	1	0.781	0.953	0.680	1.336
Age (ref: 18 years X treatment)								
16 years X Treatment	-0.212	0.217	0.962	1	0.327	0.809	0.529	1.236
17 years X Treatment	0.099	0.245	0.163	1	0.687	1.571	0.683	1.783
Treatment X Sex (ref: RAPT X female)	0.376	0.190	3.939	1	0.047	1.457	1.005	2.112
Sex X Age (ref: 18 years)								
16 years X Sex	-0.089	0.217	0.170	1	0.680	0.915	0.598	1.398
17 years X Sex	0.515	0.251	4.212	1	0.040	1.674	1.023	2.739

Given the statistically significant sex by treatment interaction, estimated mean crash rates from the overall Poisson regression model effects were calculated for sex and treatment (Table 18). These calculations indicated that exposure to RAPT for males may be associated with lower crash rates relative to males in the comparison group. However, the means suggested females exposed to RAPT may have had higher crash rates relative to the females in the comparison group. To test these findings, Poisson regression analyses were conducted for males and females separately.

Table 18. Estimated Total Crashes per Driver from Poisson Regression Model

Treatment Group	Sex	*M*	*SE*
Comparison	Male	0.1063	0.0092
	Female	0.0747	0.0086
RAPT	Male	0.0811	0.0081
	Female	0.0837	0.0086

For males, the regression analysis produced a significant main effect of treatment (Table 19), χ^2 (1, n = 2743) = 5.517, p =.019, with RAPT-trained males showing an approximately 23.7% lower crash rate relative to the male comparison group. For females, the RAPT group had an estimated 10.7% higher crash rate than the comparison group, but the main effect of treatment for females was not statistically significant (Table 20), χ^2 (1, n = 2447) = 0.553 p = .457.

Table 19. Males Poisson Regression Model Parameter Estimates for Total Crashes

			Hypothesis Test				95% CI Exp(*B*)	
Parameter	*B*	*SE*	Wald χ^2	*df*	*p*	Exp(*B*)	*LL*	*UL*
Intercept	-2.515	0.0953	696.048	1	0.000	0.081	0.067	0.097
Treatment	0.295	0.1255	5.517	1	0.019	1.343	1.050	1.718

Table 20. Females Poisson Regression Model Parameter Estimates for Total Crashes

			Hypothesis Test				95% CI Exp(*B*)	
Parameter	*B*	*SE*	Wald χ^2	*df*	*p*	Exp(*B*)	*LL*	*UL*
Intercept	-2.426	0.0945	659.123	1	0.000	0.088	0.073	0.106
Treatment	-0.104	0.1399	0.553	1	0.457	0.901	0.685	1.185

Time to First Crash

Researchers used Cox regression analysis (survival analysis) to evaluate the number of weeks after licensure at which each studied driver had their first crash (time to first crash). Prior to conducting this survival analysis, researchers tested the proportionality of hazards assumption by assessing the potential for a significant time by covariate/treatment interaction. No violations of the assumptions that would have necessitated models containing interactions with time were found. As with the Poisson analysis, analyses included calculating multiple models to assess for the best fit. None of the overall Cox regression models (Table 21) reached statistical significance. Thus an effect of RAPT on time to first crash could not be confirmed.

Table 21. Cox Regression Models for the Number of Weeks to First Crash

	Cox Proportional Hazard Models				
	A	B	C	D	E
Treatment	$p = 0.462$	$p = 0.707$	$p = 0.332$	$p = 0.370$	$p = 0.358$
Sex	$p = 0.298$	$p = 0.456$	$p = 0.520$	$p = 0.222$	
Age	$p = 0.457$	$p = 0.075$	$p = 0.653$	$p = 0.667$	
Treatment X Sex	$p = 0.057$	$p = 0.050$	$p = 0.038$		
Treatment X Age	$p = 0.857$	$p = 0.428$			
Sex X Age	$p = 0.691$	$p = 0.178$			
Treatment X Sex X Age	$p = 0.150$				
	$\chi^2 = 17.344$	$\chi^2 = 13.500$	$\chi^2 = 7.766$	$\chi^2 = 3.189$	$\chi^2 = .845$
	$df = 11$	$df = 9$	$df = 5$	$df = 4$	$df = 1$
	$p = 0.098$	$p = 0.141$	$p = 0.170$	$p = 0.527$	$p = 0.358$

For the reader interested in a detailed depiction of the survival distributions by sex, age, and treatment, Kaplan-Meier Life Tables can be found in Appendix B (Table B3 for males and B4 for females). Figure 5 and Figure 6 provide the survival curves for males and females, respectively. As can be seen, any effect of RAPT on number of weeks to first crash was negligible. For males, Figure 5 shows a minimal departure between the two curves at approximately week 10 with the RAPT group showing slightly better survival rates from that point forward. Figure 6 shows that around week 20, the curves for females have a very slight departure with the comparison group having a minimally higher survival rate compared to the RAPT females from that point forward. These curves are included for completeness, and any patterns in them must be interpreted with extreme caution since the overall models were not statistically significant.

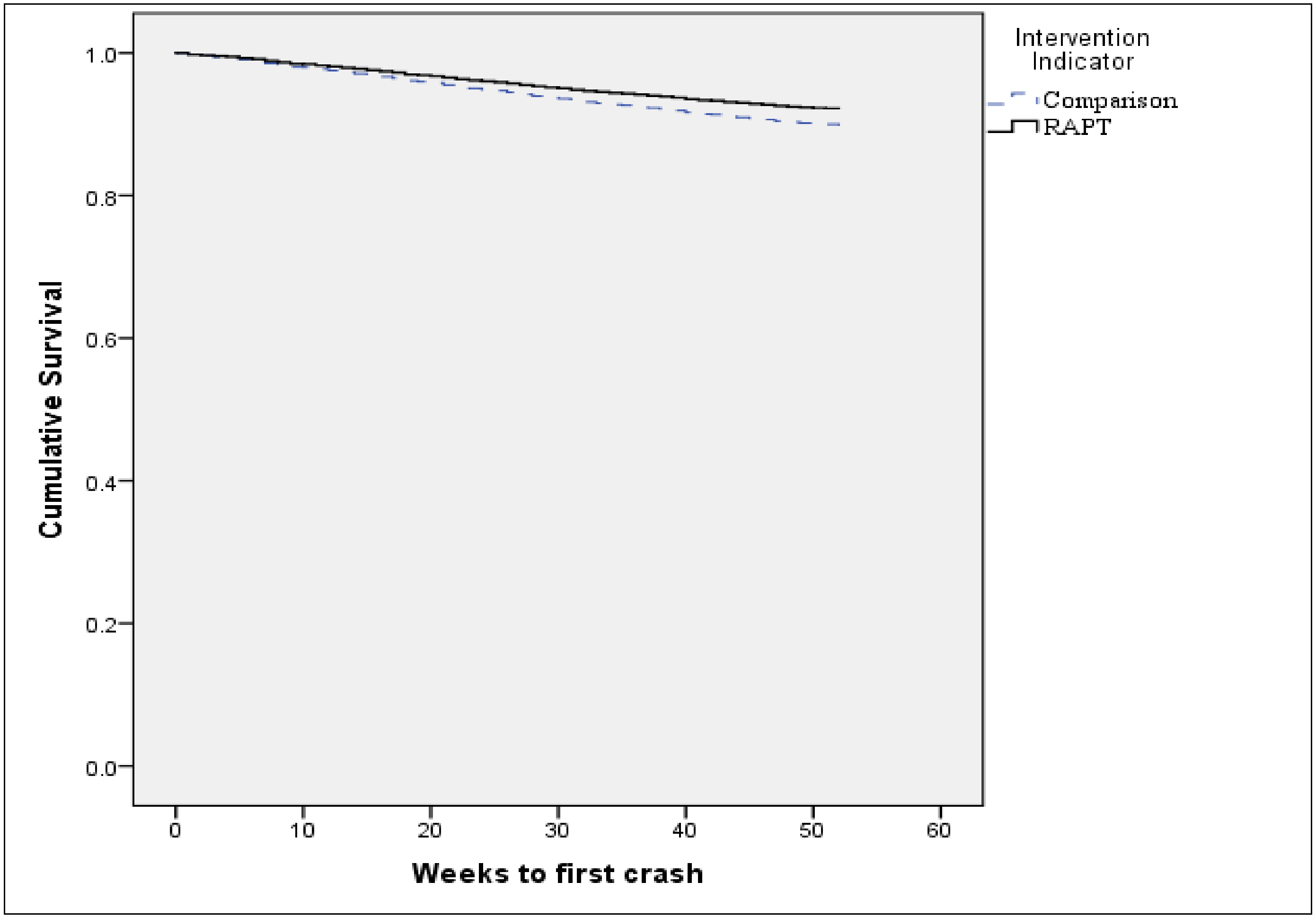


Figure 5. Survival Curve for Males by Treatment Group

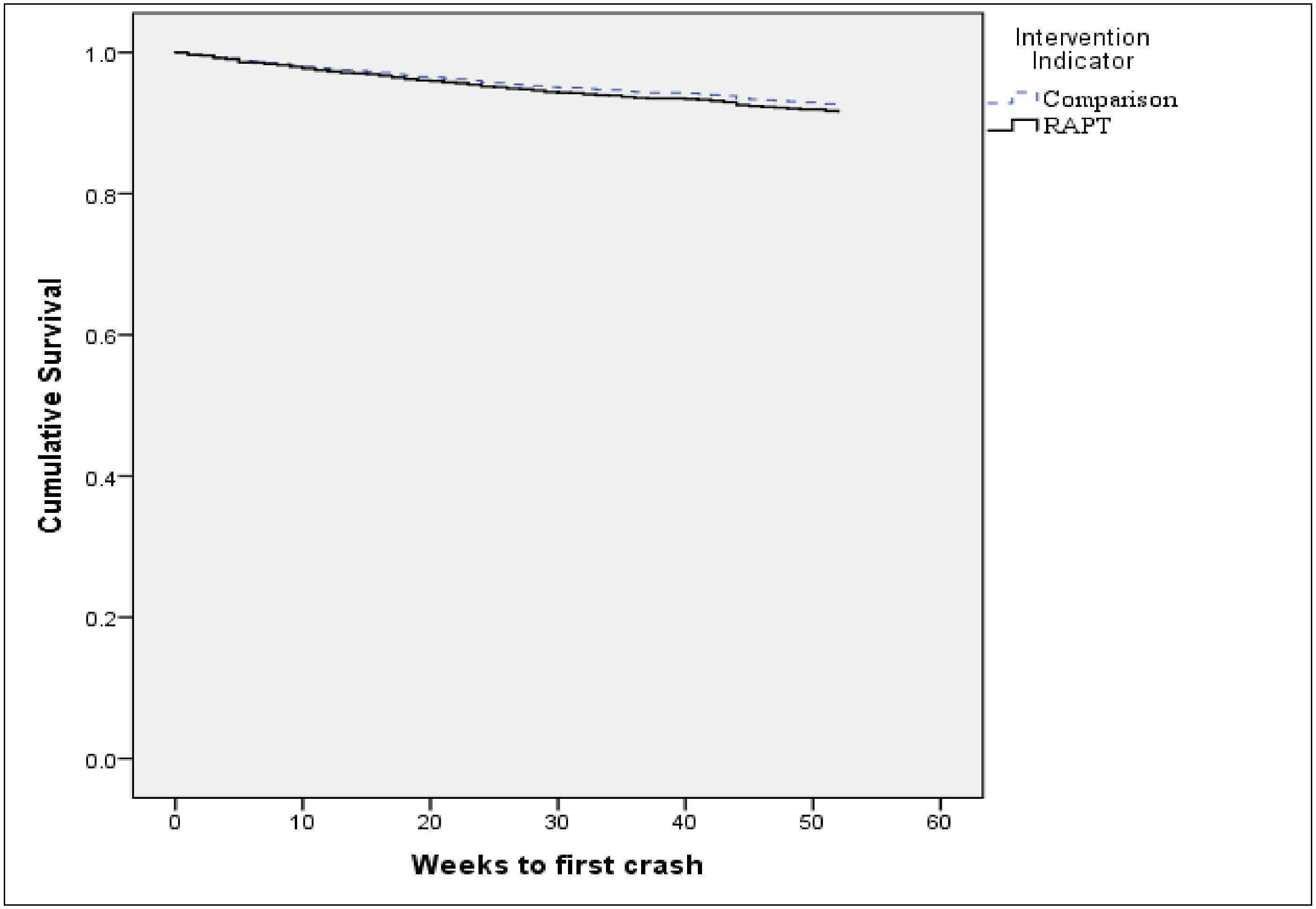


Figure 6. Survival Curve for Females by Treatment Group

RAPT Scores and Crashes

Table 22 provides the correlations of the number of defined targets correctly identified during the pre-test for all participants (RAPT and comparison conditions) with participant age, sex, and total number of crashes during the first 12 months post-licensure. RAPT pre-test scores had no relationship to crashes, but did show statistically significant correlations with age and sex. The correlation with age was negative indicating better performance for younger drivers, which is consistent with the mean pre-test scores presented earlier. The correlation with sex was negative, which indicates females performed worse than males on the pre-test. Participants in the comparison group did not complete the post-test so researchers could not calculate a change score or use any post-test measures for the entire population of participants.

Table 22. All Participants Correlations of RAPT Pre-Test, Crashes, Age, and Sex

	Age	Sex	Total crashes	Pre-Test score
Age	1.00	-0.046**	0.005	-0.119**
Sex		1.00	-0.017	-0.162**
Total crashes			1.00	0.001
Pre-test score				1.00

*Indicates significance at or below the .05 level **Indicates significance at or below the .01 level.

For the RAPT-trained participants only, correlations were calculated for the number of targets correctly identified during the pre-test, the number of targets correctly identified in the post-test, and the residualized change score between pre- and post-test with age, sex, and total number of crashes in the first 12 months post-licensure. A residualized change score is a raw change score that controls for pre-test performance by including pre-test performance as a covariate. This approach is often used when the pre-test and post-test scores are correlated. The primary advantage of this approach is that it removes the overlapping variance (Kenny, 1975). Additionally, the question that this approach answers is slightly different than the one addressed by an analysis of raw change scores. Specifically, this approach allows conclusions to be made about whether one person's performance (i.e., that of a member of group A), is expected to change more than another person's performance (i.e., that of a member of group B) (Hand & Taylor, 1987). As shown in Table 23, none of these measures of RAPT performance correlated with the number of crashes in the first 12 months post-licensure. Pre-test score was positively correlated with post-test score, but showed negative correlations with sex and age. Post-test score was negatively correlated with age and sex. These results again show that females tended to perform worse on both the pre-test and post-test, and younger participants tended to perform better on the tests than did older participants. The residualized change score was negatively correlated with sex which indicates females tended to show less improvement than males from pre-test to post-test, but a review of the means provided earlier shows this difference in change scores to be extremely small.

Table 23. RAPT Group Correlations of RAPT Scores With Crashes, Age, and Sex

	Age	Sex	Total crashes	Pre-test score	Post-test score	Residualized change score
Age	1.00	-0.076**	-0.01	-0.118**	-0.073**	-0.031
Sex		1.00	0.013	-0.175**	-0.142**	-0.087**
Total crashes			1.00	-0.015	-0.025	-0.022
Pre-test score				1.00	0.371**	0.000
Post-test score					1.00	0.925
Residualized change score						1.00

*Indicates significance at or below the .05 level **Indicates significance at or below the .01 level.

To further evaluate any potential interaction effects among sex, age, and the residualized change score, a Poisson regression was calculated for number of crashes relative to RAPT performance (Table 24). The dependent variable in this analysis was the total number of crashes in the first 12 months post-licensure. Sex and age were entered as categorical variables and residualized change score was entered as a covariate. The overall model was not significant, $\chi^2(11, n = 2{,}628) = 8.955$, $p = .626$, and neither were any of the individual covariates. Given that the overall model was not significant, researchers did not conduct an evaluation of the individual parameters. The interested reader can review the parameter estimates in Table B6 of Appendix B. The lack of a post-test and, hence, any change scores for the comparison group precludes interpretation of any association between gain scores and subsequent crash rates.

Table 24. Poisson Regression for Residualized Change Score and Crashes

	Type III		
Source	Likelihood Ratio χ^2	*df*	*p*
Intercept	3387.494	1	0.000
Sex	0.040	1	0.841
Residualized Change Score	0.498	1	0.480
Age	3.658	2	0.161
Sex X Age	2.408	2	0.300
Sex X Residualized Change Score	0.636	1	0.425
Age X Residualized Change Score	0.390	2	0.823
Sex X Age X Residualized Change Score	0.683	2	0.711

Traffic Violations

None of the analyses of violations demonstrated any association with the RAPT treatment, which is not surprising since conformance with traffic laws was not a component of the RAPT program. Because the results showed no meaningful differences in violations based on training group, the body of this report does not present the analyses and results relating to the violation data. The interested reader can find these violation results in tabular format in Appendix B.

Discussion

The objectives of this study were to update the RAPT program and to evaluate it based on the actual driving performance (i.e., crashes and violations) of a sample of teen drivers in the first year after they received their driver's license for unsupervised driving. This required the design of a study capable of recruiting a large sample of teen drivers and assigning them in an unbiased manner either to receive RAPT or to a comparison group that did not receive any training. The study achieved these objectives as the resulting dataset of over 5,000 teen drivers showed no evidence of any meaningful biases between the groups who did and did not complete RAPT. Thus, the participant assignment process employed successfully formed the basis for a comparative study of the effects of RAPT on the crash and violation records of trained and similar untrained drivers during their initial year of unsupervised driving. The resulting sample, while not a representative sample of the nation or California, includes a cross-section of young drivers.

The updated RAPT program included a data collection function that captured whether a participant clicked on the defined critical hazard area in each of nine simulated driving scenarios. The program recorded this information for a pre-test and an identical post-test for the RAPT-trained group. The comparison group only received the pre-test. Using these embedded measures, researchers found statistically significant but practically meaningless differences between the RAPT and comparison groups in terms of performance on the RAPT pre-test, providing further evidence of group equivalency. They did, however, identify substantial increases in correct performance from pre-test to post-test for the RAPT group that were consistent with previous smaller scale tests of the program. The original developers of RAPT included the pre- and post-tests as an integral part of the training program and as a rudimentary measure of change resulting from exposure to the training module. The existence of an increase in the number of correct responses after the training indicated that something was learned by the trainees but did not necessarily confirm an increase in risk perception or hazard recognition knowledge. This change in raw test scores in a positive direction did represent a necessary condition for demonstrating that the participants attended to the RAPT contents.

The vast majority (99%) of study participants could be matched confidently with their DRM data on crashes and violations without any discernable bias between the RAPT-trained and comparison groups. The crash measure was of primary interest since RAPT was focused on the identification of crash hazards and has no coverage of violation avoidance. The crash results, as represented by the Poisson regression analyses, suggested that RAPT was associated with a statistically significant and meaningful 23.7% decrease in crash rate for males. Interestingly, females exposed to RAPT showed a 10.7% higher crash rate relative to females who did not complete the training, but this difference was not statistically significant. While the three-way interaction of treatment, sex, and age was not statistically significant, it is noteworthy that 17 and 18-year-old males appear to be responsible for the crash rate decrease among the RAPT-trained male group as a whole. The apparent positive effect on male crash rates but not on the rates for females is notable because previous research (e.g., SWOV Fact Sheet, 2014) had not observed sex differences related to hazard perception training.

The full models for the Cox regression analyses of time to first crash were not statistically significant. As such, the significant terms within these models should not be interpreted.

While a differential effect of RAPT by sex was unexpected, these findings are not inconsistent with previous research focused on the teen driver. Although the precise reason for the differential response to RAPT by sex is currently unclear, factors dealing with both traffic safety and computer-based learning could have played a role.

Traffic safety research has long shown differential crash rates and driving exposure for male and female teenagers (e.g., Brar & Rickard, 2013; Masten & Foss, 2010; Vlakveld, 2011) with males driving more miles and having higher crash incidence and more severe crashes. These documented differences between the sexes lead to several plausible explanations for the observed pattern of results. Because of their greater exposure to crashes, males may have had a better opportunity to demonstrate a crash reduction associated with exposure to RAPT. Also, male teens admit to more driving at excessive speeds (e.g., Scott-Parker, Watson, King, & Hyde, 2014), which limits the time for hazard recognition and response. Since RAPT focuses on scanning behavior, it may have overcome to some extent the difficulty in detecting hazards when speeding. It must be noted that this assumes that the male drivers have generalized the static images of travelling on surface streets in RAPT to higher speed driving or driving at inappropriate speeds for conditions.

A second possible contributor to the pattern of differential RAPT results by sex might stem from the relative effectiveness of computer-based training programs for males and females and the generally different approach of the sexes to computers (e.g., Sims, Chin, Durrance, & Johnson, 2004; Christoph, Goldhammer, Zylka, Johannes, & Hartig, 2015; Cassidy & Eachus, 2002). Prior research has shown that, while females do improve from computer-based training, their improvement tends to be less than that of their male counterparts (e.g., Sanchez-Ku & Arthur Jr., 2000; Turos & Ervin, 2000). Given these sex differences, a program such as RAPT might be expected to work better for male than for female teenagers. Additional research would be needed to confirm or refute this theory.

It is also important to note the entire RAPT approach focuses on improved scanning behavior as a means of identifying hazards and reducing crashes. Instructions to RAPT participants involved clicking on areas in the presented pictures where they would be looking if driving. The feedback during the training module emphasizes where participants did (or should have) clicked to indicate they were looking at a potential hazard. This may have improved hazard perception or risk awareness which could result in a crash reduction, or it may have served to increase scanning behavior while driving, which also could have produced a safety benefit. In either case, scores on the rudimentary and identical pre- and post-tests were not associated with crashes which means they may not be a valid measure of risk perception or hazard recognition. They certainly have never been previously assessed for their predictive ability, nor has anyone ever suggested the scores would be related to crash rates. In addition, such computer-based assessments that rely on mouse clicks on critical displayed scene areas as in the current study may not be valid measures of risk perception or hazard recognition since experienced drivers often perform about the same or worse on these tests as novice drivers despite the better hazard recognition capabilities of the experienced drivers as measured by eye-tracking on the road or in a simulator (e.g., Vlakveld, 2011). In short, while the current study suggests that completing RAPT was associated with lower crash rates for male teens, the precise mechanism producing this association is not yet clear. More research is therefore necessary to better understand how RAPT works, and for whom.

Conclusions

Overall, the results of this study provide perhaps the first encouraging evidence that brief, computer-based training interventions can have a positive influence on driving safety for newly licensed teen drivers even if only for males. Given the size of the sample included in this study, the single State venue, the fact that the RAPT implementation studied used only nine selected scenarios, and the experimental limitations of the study discussed below, the reader must exercise caution when generalizing these findings. Nevertheless, the potential importance of the existence of some positive crash-based results for guiding future research and development with respect to the driver training process cannot be overlooked. Further research is needed to clarify the uncertainties arising from this study, particularly related to the lack of effectiveness, and even potential detrimental effects, of RAPT on female crash rates. Once a better understanding of the effects of RAPT is achieved, it would be useful to assess how best to employ hazard perception training using a program such as RAPT in the driver training process.

Limitations

The reader should bear in mind the limitations of this study when interpreting its findings. Some of these limitations are inherent in research of this type while others are unique to the approach selected for this particular study. First, while a sample of over 5,000 newly licensed drivers enlisted at precisely the same point in their driving careers might appear robust, it is, in fact, just adequate for a study of this type. This is evidenced by the relatively small number of crashes across the entire sample. A larger crash sample size might have shed light on the questions of whether the pattern of results reported herein was accurate or simply an artifact of the sample size studied, and whether RAPT had a differential effect by crash severity.

A second limitation in this study was self-imposed by the researchers. Potentially valuable information for interpreting the pattern of results might have been produced if the comparison group had been given the same post-test as the RAPT-trained participants. The inclusion of a placebo training module for the comparison group was considered and rejected because of time and cost considerations. As a result, the study treated the two groups of participants totally equivalently only up to the completion of the pre-test. Thereafter, the comparison group was dismissed and the experimental group continued through the training and post-test. This differential experience for the RAPT group could have led to effects unrelated to the training that could not be assessed with the chosen study design.

Another concern associated with the RAPT task is that the underlying mechanism associated with positive outcomes among teen males is simply unknown. Available measures that would normally be used to establish predictive validity—such as gain scores from the pre-test to the post-test—show no relationship to the main outcome of interest, namely crashes. This is not surprising, however, since the pre-tests and post-tests were never designed to be predictive of crashes; rather they were included as rudimentary measures of whether participants were attending to the training materials. Given this design, the post-test scores may simply be a measure of a subject's short-term spatial recognition memory and not a measure of actual learning.

In studies of this type, refusals or drop-outs can produce significant bias. In the present study, the estimated participation rate was 77.9%, which is quite high, but there was substantial variability of participation rates across DMV offices, ranging from 50% to 97%. Despite this variability, the analyses of the equivalency of the RAPT and comparison groups indicated that no detectable biases existed between the two groups that entered the study. It also must be noted that nothing is known about the young drivers who were eligible to participate but did not. Additionally, the drivers who participated in this study may not be representative of the population of California teen drivers based on socio-economic status and ethnicity. For these reasons, it is important to exercise caution when generalizing these results.

Finally, the reader must remember the results obtained arose from a single research study and not an operational implementation of RAPT. Participants received RAPT training in a convenience sample of six California DMV licensing offices immediately after passing the drive test. It was also clear to the participants that they were in a research study and that RAPT was not part of their licensing requirements. The study protocol included no notion of passing or failing RAPT. The participants received payment for taking part in the research rather than paying for RAPT as part of driver training or the licensing process. These circumstances may have affected the attention of the participants to the material in some positive or negative manner that cannot be discerned from the available information.

References

Brar, S., & Rickard, D. (2013). *Teen and senior drivers* (Report No. CAL-DMV-RSS-13-240). Sacramento, CA: California Department of Motor Vehicles.

Cassidy, S., & Eachus, P. (2002). Developing the computer user self-efficacy (CUSE) scale: Invesigating the relationship between computer self-efficacy, gender and experience with computers. *Journal of Educational Computing Research, 26*(2), 133-153.

Chapman, P., Underwood, G., & Roberts, K. (2002). Visual search patterns in trained and untrained novice drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 5*(2), 157-167.

Christoph, G., Goldhammer, F., Zylka, Johannes, & Hartig, J. (2015). Adolescents' computer performance: The role of self-concept and motivational aspects. *Computers & Education*, *81*(2015), 1-12.

Fisher, D. L., Laurie, N. E., Glaser, R., Connerney, K., Pollatsek, A., Duffy, S. A., et al. (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*, 287-302.

Fisher, D. L., Pradhan, A. K., Pollatsek, A., & Knodler, M. A. J. (2007). Empirical evaluation of hazard anticipation behaviors in the field and on a driving simulator using an eye tracker. *Proceedings of the 86th Transportation Research Board Annual Meeting* (CD-ROM), TRB, National Research Council, Washington, DC.

Fisher, D. L., Thomas, F. D., Pradhan, A. K., Pollatsek, A., Blomberg, R. D. & Reagan, I. (2010, January). Development and evaluation of a PC-based attention maintenance training program (Report No. DOT HS 811 252). Washington, DC: National Highway Traffic Safety Administration.

Hand, D. J., & Taylor, C. C. (1987). *Multivariate analysis of variance and repeated measures: A practical guide for behavioural scientists.* London: Chapman & Hall.

Insurance Institute for Highway Safety (n.d.). Teenagers: Driving carries extra risk for them. Retrieved on August 25, 2015 from the IIHS website at www.iihs.org/iihs/topics/t/teenagers/fatalityfacts/teenagers.

Kenny, D. A. (1975). A quasi-experimental approach to assessing treatment effects in the nonequivalent control group design. *Psychological Bulletin, 82*(3)*,* 345-362.

Masten, S. V., & Hagge, R. A. (2004). Evaluation of California's graduated driver licensing program. *Journal of Safety Research*, *35*(5), 523-535.

Masten, S. V., & Foss, R. D. (2010). Long-term effect of the North Carolina graduated driver licensingsystem on licensed driver crash incidence: A 5-year survival analysis. *Accident Analysis & Prevention, 42*:1647–1652.

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

Pollatsek, A., Narayanaan, V., Pradhan, A.K., & Fisher, D.L. (2006). Using eye movements to evaluate a PC-based risk awareness and perception training program on a driving simulator. *Human Factors, 48*, 447–464.

Pradhan, A. K., Hammel, K. R., DeRamus, R., Pollatsek, A., Noyce, D. A., & Fisher, D. L. (2005). The use of eye movements to evaluate the effects of driver age on risk perception in an advanced driving simulator. *Human Factors, 47*, 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, 62*, 657-673.

Regan, M. A., Triggs, T. J., & Godley, S. T. (2000). *Simulator-based evaluation of the DriveSmart novice driver CD-ROM training product.* Paper presented at the Proceedings of the Australasian 2000 Road Safety Research, Policing, and Education Conference.

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

Sanchez-Ku, M., & Arthur Jr., W. (2000). A dyadic protocol for training complex skills: A replication using female participants. *Human Factors, 42*(3), 512-520.

Scott-Parker, B., Watson, B., King, M., & Hyde, M. (2014). “I drove after drinking alcohol” and other risky driving behaviours reported by young novice drivers. *Accident Analysis and Prevention, 70*(2014), 65-73

Sims, V., Chin, M., Durrance, L., & Johnson, C. (2004). Mental models of computers as a function of sex and perceived trust. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting*, New Orleans, LA, September 20-24, 2004.

SWOV (2014). S*WOV Fact Sheet: Hazard perception and how to test it*. The Hague, The Netherlands: SWOV Institute for Road Safety Research.

Turos, J., & Ervin, A. (2000). Training and gender differences on a web-based mental rotation task. *The Penn State Behrend Psychology Journal, 4*(2), 3-12.

Vlakveld, W. (2011). *Hazard anticipation of young novice drivers: Assessing and enhancing the capabilities of young novice drivers to anticipate latent hazards in road and traffic situations*. Geelong, Australia: Transport Accident Commission.

Appendix A

Select Study Forms and Materials

Information about the Young Driver Safety Study

Background

Newly licensed teen drivers have a crash rate during their first month of unsupervised driving that is up to five times higher than the crash rate of teen drivers who have had their license six months or more. The crash rate differences are even more pronounced when newly licensed teen drivers are compared to older and more experienced drivers. In response to these crash differences, research has been focused at reducing young driver crash rates by using innovative computer-based programs designed to increase safe driving behaviors.

Study Objectives

The objective of this study funded by the National Highway Traffic Safety Administration (NHTSA) is to amass a large sample of newly licensed teen drivers (18 years of age or younger) who work through an innovative computer program immediately after they pass the driver licensing road exam. A subsequent NHTSA-sponsored study will evaluate the program and determine if it reduces the crash and violation rates of young drivers.

Contract Information

The work is being performed by Dunlap and Associates, Inc. under Task Order 0001 of Contract DTNH22-06-D-00032 from the National Highway Traffic Safety Administration.

No California DMV or State money is involved in this study.

Dunlap and Associates, Inc. 110 Lenox Avenue, Stamford, Connecticut 06906-2300 (203) 323-8464
FAX (203) 964-0799

Figure A1. **Exit Flyer.**

Participant Count and Recruitment Flyer Tally Form

Office Location:

Date: ________________________

Proctor Name: ________________________

a. Total number of Recruiting Flyers given to Technicians/Examiners at the start of the day (normally 50) __________________

Computer Number	Tally of Participants (tick marks)	Total Participants (number of tick marks for each computer)
Total		**b.**

c. Total Number of Flyers Collected from the Technicians/Examiners at the end of the day (remainder of the 50) __________________

Dunlap and Associates, Inc. 110 Lenox Avenue, Stamford, Connecticut 06906-2300 (203) 323-8464
FAX (203) 964-0799

Figure A2. **Proctor Tally Form.**

Agreement to Participate

The research study you have been asked to take part in is funded by the U.S. Department of Transportation, National Highway Traffic Safety Administration (NHTSA) and is being conducted by Dunlap and Associates, Inc. as a contractor for NHTSA. **This research study is not being funded by the State of California or the California DMV.**

You understand that:

- You are voluntarily participating in a safety research project for which you must provide your driver's license number so that your driving record can be included in the research statistics.
- Your information will be kept confidential.
- As payment for participating, a check for $30 will be sent to the name and address you enter and will take 2-4 weeks to arrive.
- You can stop and leave this office at any time for any reason.

☐ Click here to acknowledge you have read the above and agree to proceed with the study

Figure A3. **Agreement to Participate Computer Screen.**

Appendix B

Supplemental Statistical Analyses

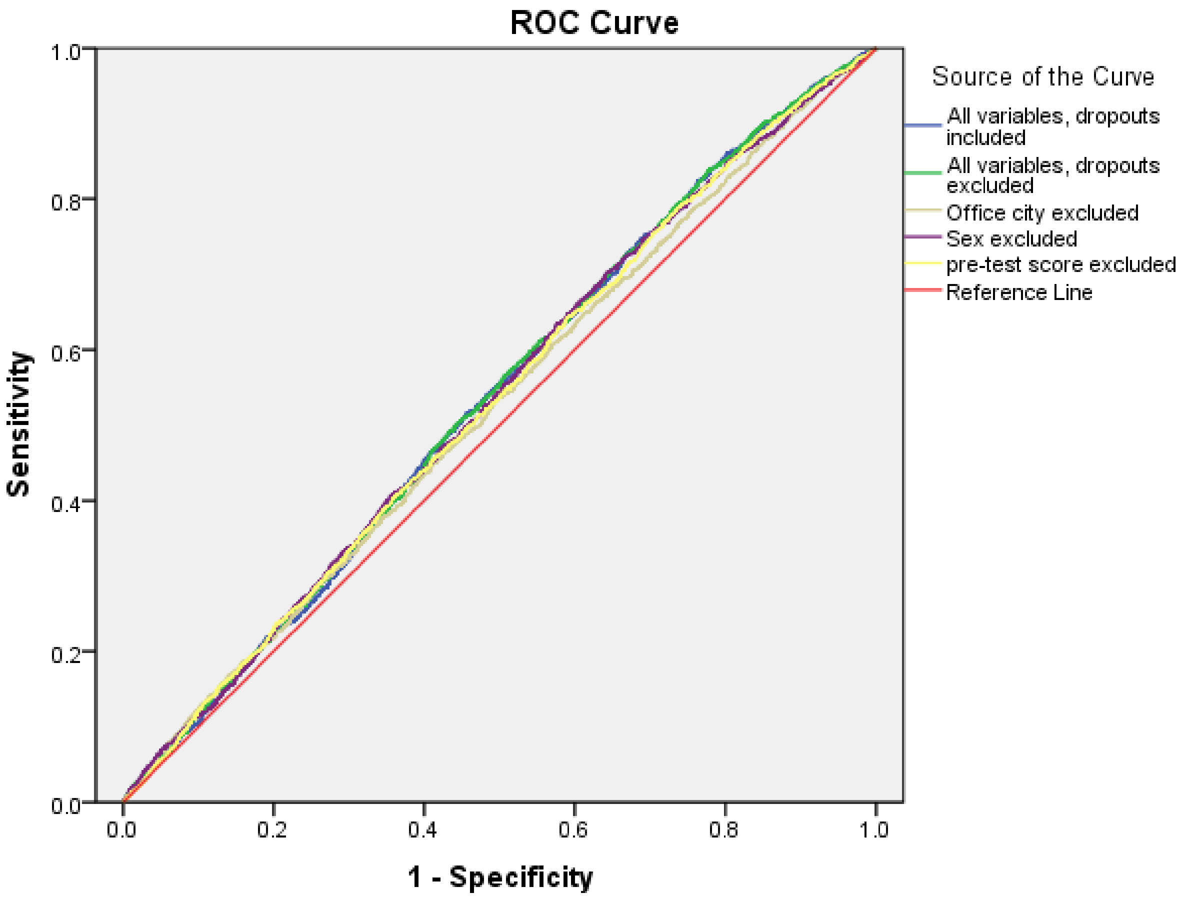


Figure B1. **ROC curves for each of the five models tests.**

Table B1. Area Under the Curve (AUC) for Each of the Regression Models

Regression Model	AUC
All variables, drop-outs excluded	0.536
All variables, drop-outs included	0.537
Office City excluded	0.524
Sex excluded	0.534
Pre-test scores excluded	0.532

Table B2. PDO and Injury Crashes During the First 12 Months Post-Licensure

	Placebo		RAPT	
	Male	Female	Male	Female
PDO crashes	111	63	86	79
Injury and fatal crashes	28	24	21	27
Total crashes	139	87	107	106

Table B3. Kaplan Meier life tables for the number of weeks to first crash for males only.

Age	Treatment Condition	Interval (in weeks)	Number Entering Interval	Number of Terminal Events	Proportion Surviving	Cumulative Proportion Surviving	Hazard Rate
16	**Comparison**	0-12	412	5	0.9879	0.9879	0.0009
		13-25	407	12	0.9705	0.9587	0.0023
		26-38	395	6	0.9848	0.9442	0.0012
		39-52	389	8	0.9794	0.9248	0.0016
	RAPT	0-12	386	7	0.9819	0.9819	0.0014
		13-25	379	15	0.9604	0.9430	0.0031
		26-38	364	4	0.9890	0.9326	0.0008
		39-52	360	9	0.9750	0.9093	0.0019
17	**Comparison**	0-12	333	10	0.9700	0.9700	0.0023
		13-25	323	10	0.9690	0.9399	0.0024
		26-38	313	10	0.9681	0.9099	0.0025
		39-52	303	10	0.9670	0.8799	0.0026
	RAPT	0-12	311	5	0.9839	0.9839	0.0012
		13-25	306	5	0.9837	0.9678	0.0013
		26-38	301	7	0.9767	0.9453	0.0018
		39-52	294	5	0.9830	0.9293	0.0013
18	**Comparison**	0-12	637	16	0.9749	0.9749	0.0020
		13-25	621	20	0.9678	0.9435	0.0025
		26-38	601	19	0.9684	0.9137	0.0025
		39-52	582	12	0.9794	0.8948	0.0016
	RAPT	0-12	664	17	0.9744	0.9744	0.0020
		13-25	647	7	0.9892	0.9639	0.0008
		26-38	640	17	0.9734	0.9383	0.0021
		39-52	623	8	0.9872	0.9262	0.0010

Table B4. Kaplan Meier Life Tables for the Number of Weeks to First Crash for Females Only

Age	Treatment Condition	Interval (in weeks)	Number Entering Interval	Number of Terminal Events	Proportion Surviving	Cumulative Proportion Surviving	Hazard Rate
16	**Comparison**	0-12	366	11	0.9699	0.9699	0.0023
		13-25	355	7	0.9803	0.9508	0.0015
		26-38	348	5	0.9856	0.9372	0.0011
		39-52	343	9	0.9738	0.9126	0.0020
	RAPT	0-12	427	13	0.9696	0.9696	0.0024
		13-25	414	10	0.9758	0.9461	0.0019
		26-38	404	7	0.9827	0.9297	0.0013
		39-52	397	8	0.9798	0.9110	0.0016
17	**Comparison**	0-12	287	7	0.9756	0.9756	0.0019
		13-25	280	3	0.9893	0.9652	0.0008
		26-38	277	2	0.9928	0.9582	0.0006
		39-52	275	2	0.9927	0.9512	0.0006
	RAPT	0-12	320	6	0.9813	0.9813	0.0015
		13-25	314	2	0.9936	0.9750	0.0005
		26-38	312	7	0.9776	0.9531	0.0017
		39-52	305	5	0.9836	0.9375	0.0013
18	**Comparison**	0-12	527	13	0.9753	0.9753	0.0019
		13-25	514	12	0.9767	0.9526	0.0018
		26-38	502	7	0.9861	0.9393	0.0011
		39-52	495	8	0.9838	0.9241	0.0013
	RAPT	0-12	520	12	0.9769	0.9769	0.0018
		13-25	508	17	0.9665	0.9442	0.0026
		26-38	491	9	0.9817	0.9269	0.0014
		39-52	482	9	0.9813	0.9096	0.0014

Table B5. Cox Proportional Hazard Regression Model for Weeks to First Crash Including All Covariates and Interactions

Source	*B*	*SE*	Wald χ^2	*df*	*p*	Exp(*B*)	95.0% CI for Exp(*B*)	
							LL	*UL*
Treatment	-.157	.214	0.542	1	.462	0.854	0.562	1.299
Sex	-.212	.204	1.082	1	.298	0.809	0.542	1.207
Age (ref: 18 years)			1.565	2	.457			
16 years	.011	.217	0.003	1	.959	1.011	0.661	1.546
17 years	-.294	.258	1.292	1	.256	0.746	0.449	1.237
Treatment X Sex	.540	.284	3.616	1	.057	1.717	0.984	2.997
Treatment X Age (ref: 18 years)			0.309	2	.857			
Treatment X Age (16 years)	.108	.320	0.114	1	.735	1.114	0.595	2.087
Treatment X Age (17years)	-.115	.397	0.084	1	.772	0.891	0.409	1.941
Sex X Age (ref: 18 years)			0.738	2	.691			
Sex X Age (16 years)	.230	.308	0.557	1	.455	1.259	0.688	2.304
Sex X Age (17 years)	.245	.364	0.451	1	.502	1.277	0.626	2.607
Treatment X Sex X Age (ref: 18 years)			3.790	2	.150			
Treatment X Sex X Age (16 years)	-.684	.444	2.380	1	.123	0.504	0.211	1.203
Treatment X Sex X Age (17 years)	.289	.513	0.316	1	.574	1.334	0.488	3.649

Table B6. Poisson Regression Model Parameter Estimates for Total Crashes Relative to Residualized Change Scores

Parameter	*B*	*SE*	Hypothesis Test Wald χ^2	*df*	*p*	Exp(*B*)	95% Wald CI for Exp(*B*) *LL*	*UL*
Intercept	-2.281	0.142	256.465	1	0.000	0.102	0.077	0.135
Sex (ref: female)	-0.321	0.204	2.474	1	0.116	0.726	0.486	1.082
Residualized Change Score	-0.017	0.063	0.070	1	0.791	0.984	0.870	1.112
Age (ref: 18 years)								
16 years	-0.079	0.216	0.134	1	0.715	0.924	0.605	1.411
17 years	-0.552	0.278	3.939	1	0.047	0.576	0.334	0.993
Sex X Residualized Change Score (ref: female)	-0.048	0.096	0.251	1	0.616	0.953	0.789	1.150
Sex X Age (ref: female X 18 years)								
Sex X Age (16 years)	0.328	0.311	1.115	1	0.291	1.389	0.755	2.554
Sex X Age (17 years)	0.545	0.380	2.054	1	0.152	1.724	0.818	3.632
Residualized Change Score X Age (ref: 18 years)								
Residualized Change Score X Age (16 years)	-0.042	0.104	0.164	1	0.685	0.959	0.782	1.176
Residualized Change Score X Age (17 years)	0.101	0.138	0.538	1	0.463	1.107	0.844	1.451
Residualized Change Score X Sex X Age (ref: female X 18 years)								
Residualized Change Score X Sex X Age (16 years)	0.066	0.158	0.175	1	0.676	1.068	0.783	1.457
Residualized Change Score X Sex X Age (17 years)	-0.102	0.187	0.294	1	0.588	0.903	0.626	1.304

Table B7. Negative Binomial Models for Violations by Sex, Age, and Treatment

	Models				
	A	B	C	D	E
Intercept	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$
Treatment	$p = 0.691$	$p = 0.653$	$p = 0.946$	$p = 0.744$	$p = 0.638$
Sex	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$	
Age	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$	
Treatment X Sex	$p = 0.194$	$p = 0.228$	$p = 0.238$		
Treatment X Age	$p = 0.155$	$p = 0.162$			
Sex X Age	$p = 0.987$	$p = 0.970$			
Treatment X Sex X Age	$p = 0.606$				
	$\chi2 = 68.086$	$\chi2 = 135.176$	$\chi2 = 131.463$	$\chi2 = 130.069$	$\chi2 = 0.638$
	$df = 11$	$df = 9$	$df = 5$	$df = 4$	$df = 1$
	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.000$	$p = 0.222$

Table B8. Negative Binomial Test for Violations by Sex, Age, and Treatment

	Type III		
Source	Likelihood Ratio χ^2	*df*	*p*
Intercept	2021.805	1	0.000
Treatment	0.158	1	0.691
Sex	46.593	1	0.000
Age	54.475	2	0.000
Treatment X Sex	1.686	1	0.194
Treatment X Age	3.731	2	0.155
Sex X Age	0.026	2	0.987
Treatment X Sex X Age	1.003	2	0.606

Table B9. Negative Binomial Regression Parameter Estimates for Number of Violations

Parameter	*B*	*SE*	Hypothesis Test Wald χ^2	*df*	*p*	Exp(*B*)	95% Wald CI for Exp(*B*) *LL*	*UL*
Intercept	-1.859	.120	242.298	1	.000	0.156	0.123	0.197
Treatment (ref: RAPT)	-.117	.173	0.461	1	.497	0.889	0.634	1.248
Sex (ref: female)	.565	.146	15.014	1	.000	1.760	1.322	2.342
Age (ref: 18 years)								
16 years	-.534	.206	6.745	1	.009	0.586	0.392	0.877
17 years	-.651	.236	7.584	1	.006	0.522	0.328	0.829
Treatment X Sex	.114	.210	0.294	1	.588	1.121	0.742	1.692
Treatment X Age (ref: 18 years)								
Treatment X 16 years	-.257	.323	0.633	1	.426	0.774	0.411	1.456
Treatment X 17 years	.369	.329	1.260	1	.262	1.447	0.759	2.757
Sex X Age (ref: 18 years)								
Sex X 16 years	-.216	.268	0.648	1	.421	0.806	0.477	1.363
Sex X 17 years	-.010	.298	0.001	1	.972	0.990	0.552	1.775
Sex X Treatment X Age (ref: 18 years)								
Sex X Treatment X 16 years	.377	.400	0.889	1	.346	1.458	0.666	3.194
Sex X Treatment X 17 years	-.029	.408	0.005	1	.944	0.972	0.437	2.163

Table B10. Cox Regression Analysis for Weeks to First Violation

	Cox Proportional Hazard Models				
	A	B	C	D	E
Treatment	*p* = .631	*p* = .335	*p* = .332	*p* = .857	*p* = .929
Sex	*p* = .001	*p* = .000	*p* = .520	*p* = .000	
Age	*p* = .009	*p* = .000	*p* = .653	*p* = .000	
Treatment X Sex	*p* = .671	*p* = .273	*p* = .038		
Treatment X Age	*p* = .368	*p* = .434			
Sex X Age	*p* = .878	*p* = .928			
Treatment X Sex X Age	*p* = .621				
	$\chi 2$ = 116.621 *df* = 11 *p* = .000	$\chi 2$ = 116.441 *df* = 9 *p* = .000	$\chi 2$ = 17.344 *df* = 3 *p* = .000	$\chi 2$ = 111.355 *df* = 3 *p* = .000	$\chi 2$ = .008 *df* = 1 *p* = .929

Table B11. Cox Regression Model for Weeks to First Violation

Source	*B*	*SE*	Wald χ^2	*df*	*p*	Exp(*B*)	95.0% CI for Exp(*B*) *LL*	*UL*
Treatment	-.084	.174	0.231	1	.631	0.920	0.654	1.294
Sex	.514	.148	12.114	1	.001	0.533	0.332	.855
Age (ref: 18 years)			9.383	2	.009			
16 years	-.475	.206	5.319	1	.021	0.622	0.415	.931
17 years	-.629	.241	6.801	1	.009	0.533	0.332	.855
Treatment X Sex	.090	.212	0.181	1	.671	1.094	0.722	1.658
Treatment X Age (ref: 18 years)			2.000	2	.368			
Treatment X Age (16 years)	-.368	.329	1.250	1	.264	0.692	0.363	1.319
Treatment X Age (17years)	.183	.342	0.284	1	.594	1.200	0.613	2.349
Sex X Age (ref: 18 years)			0.259	2	.878			
Sex X Age (16 years)	-.103	.265	0.151	1	.698	0.902	0.536	1.518
Sex X Age (17 years)	.064	.302	0.044	1	.833	1.066	0.590	1.925
Treatment X Sex X Age (ref: 18 years)			0.954	2	.621			
Treatment X Sex X Age (16 years)	.381	.403	0.992	1	.345	1.464	0.664	3.228
Treatment X Sex X Age (17 years)	.004	.421	0.000	1	.992	1.004	0.440	2.293

χ^2 (11, n = 5,190) = 116.621, p = .000

Table B12. Correlations of RAPT Scores, Violations, Age, and Sex

	Age	Sex	Total violations	Pre-test score	Post-test score	Residualized change score
Age	1.000	0.076**	.113**	-0.118**	-0.073**	-0.031
Sex		1.000	-0.098**	-0.175**	-0.142**	-0.087**
Total violations			1.000	-0.048*	-0.066**	-0.035
Pre-test score				1.000	0.371**	0.000
Post-test score					1.000	0.925
Residualized change score						1.000

* indicates significance at or below the .05 level ** indicates significance at or below the .01 level.

Table B13. Negative Binomial Regression Test of Violations by Residualized Change Score

	Type III		
Source	Likelihood Ratio χ^2	*df*	*p*
Intercept	1689.196	1	.000
Sex	16.826	1	.000
Residualized Change Score	2.938	1	.087
Age	29.299	2	.000
Sex X Residualized Change Score	1.912	1	.167
Sex X Age	0.563	2	.755
Age X Residualized Change Score	0.087	2	.957
Sex X Age X Residualized Change Score	0.453	2	.797

Table B14. Negative Binomial Regression Parameter Estimates for Sex, Age, Residualized Change Score, and Violations

			Hypothesis Test				95% Wald CI for Exp(*B*)	
Parameter	*B*	*SE*	Wald χ^2	*df*	*p*	Exp(*B*)	*LL*	*UL*
Intercept	-1.907	.1281	221.800	1	.000	0.149	0.116	0.191
Sex (ref: female)	.610	.1542	15.637	1	.000	1.840	1.360	2.489
Residualized Change Score	-.082	.0524	2.437	1	.119	0.921	0.831	1.021
Age (ref: 18 years)								
16 years	-.522	.2178	5.745	1	.017	0.593	0.387	0.909
17 years	-.617	.2492	6.119	1	.013	0.540	0.331	0.880
Sex X Residualized Change Score	.038	.0681	0.309	1	.579	1.039	0.909	1.187
Sex X Age (ref: 18 years)								
Sex X 16 years	-.210	.2806	0.561	1	.454	0.811	0.468	1.405
Sex X 17 years	-.045	.3109	0.021	1	.885	0.956	0.520	1.759
Residualized Change score X Age (ref: 18 years)								
Residualized Change Score X 16 years	-.020	.1027	0.039	1	.844	0.980	0.801	1.199
Residualized Change Score X 17 years	-.031	.1054	0.088	1	.766	0.969	0.788	1.192
Residualized Change Score X Sex X Age (ref: 18 years)								
Residualized Change Score X Sex X 16 years	.082	.1438	0.328	1	.567	1.086	0.819	1.439
Residualized Change Score X Sex X 17 years	.067	.1421	0.224	1	.636	1.070	0.810	1.413

DOT HS 812 235
January 2016

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

NHTSA
www.nhtsa.gov

12029-010716-v2