EFFECTS OF INACCURATE GAZE BEHAVIOR ON YOUNG DRIVERS' HAZARD ANTICIPATION

Sarah Yahoodik[1], Nathan Hatfield[2], Yusuke Yamani[1], Siby Samuel[3]
[1]Old Dominion University, Norfolk, VA, USA
[2]Design Interactive, Inc., Williamsburg, VA, USA
[3]University of Waterloo, Waterloo, Ontario, Canada
Email: syaho001@odu.edu

Summary: A previous study (Yamani et al., 2018) demonstrated that the administration of expert eye movement videos following hazard anticipation training can improve the proportion of latent hazards anticipated by young drivers compared to control conditions. The current driving simulator study sought to examine whether the improvements observed in the previous study were merely due to drivers' exposure to videos of the simulated driving scenarios with expert eye movement overlays immediately prior to evaluation, or whether modeling the accuracy of eye movement behavior can lead participants to internalize hazard anticipation skills more effectively. In a between-subject design, 36 drivers (18-21 years) were assigned to one of three experimental conditions – training only, training plus expert eye movements or training plus novice eye movements. All participants navigated four unique driving scenarios, each with their eye movements tracked and recorded. Analyses of the eye movement data showed that young drivers who saw the expert eye movement (accurate) videos immediately following training anticipated a substantially greater proportion of latent hazards compared to the young drivers that saw novice eye movement (inaccurate) videos following training. The data provide some evidence that drivers were able to successfully map and incorporate correct hazard anticipation glance behavior into their mental models. The findings present some implications for the design and evaluation of eye movement-based training interventions.

INTRODUCTION

While the number of young drivers involved in fatal crashes has decreased over the past decade, the fatal crash rate per mile driven for 16-19 year olds is still almost three times the crash rate of all other drivers (IIHS, 2017; NHTSA, 2018). Research has suggested that factors such as lack of sufficient attention and visual search behavior contribute heavily to novice driver crashes (McKnight & McKnight, 2003). Newly licensed young drivers (in the 16-17 age range) are most at risk of failing to recognize hidden hazards out of all age groups, but young drivers (19-29 years old) are still significantly worse at recognizing latent hazards than older adults (60-75 years old) (Pradhan et al., 2005). Many training programs geared towards young novice drivers have focused on addressing hazard anticipation errors via a variety of modalities, such as expert narration, computer-based trainings and desktop-simulator programs, which have demonstrated both short-term and long-term success in improving hazard anticipation behavior (Unverricht, Samuel & Yamani, 2018; McDonald, Goodwin, Pradhan, Romoser, & Williams, 2015).

One such training program is the Risk Awareness and Perception Training (RAPT). RAPT is designed to train and improve the latent hazard anticipation behavior of young drivers on the road (Pradhan, Pollatsek, Knodler, & Fisher, 2009). RAPT trainees are shown examples of

driving scenarios that involve latent hazards where either the potential collision risk is hidden from view or where a lead vehicle may brake suddenly due to vehicle or pedestrian behavior. RAPT has been evaluated and proven effective in improving latent hazard anticipation among young novice drivers in the simulator and on-road both immediately after training and for up to six months after training (Taylor et al., 2011). When novice drivers were reevaluated six months after initially being trained, the RAPT group correctly anticipated latent threats 61.9% of the time versus the control group which only did such 37.7% of the time.

Despite the demonstrated ability of RAPT and its variants to substantially improve the hazard anticipation behavior of young novice drivers, the anticipation performance of novice drivers does not approach that of experienced drivers (Unverricht et al., 2018). One potential way to further improve on RAPT may be to demonstrate to novice drivers how experienced drivers scan a dynamic driving scenario. Because RAPT consists solely of static images, trainees may not have been able to transfer the skills they learned to a dynamic driving scenario.

In a previous study, Yamani et al. (2018) examined the effects of showing young drivers how experts distribute their visual attention when navigating a dynamic driving scenario. Young drivers completed RAPT and drove four unique simulated driving scenarios with one hazard anticipation scenario each. Trainees' anticipatory glances towards the location of a latent hazard were evaluated. One group only completed the RAPT module, while another group of participants also watched a 10-minute video of the simulated driving scenarios that displayed the eye movements via crosshair of an "expert" driver who perfectly anticipated the latent hazards in the scenario. A third group only watched the 10-minute video of the expert eye glances before completing the hazard anticipation task. The results indicated that completing RAPT followed by watching the glance video led to better hazard anticipation performance than simply completing RAPT or watching the video. However, it remains unexplored whether this effect was due to participants successfully incorporating proper visual search behavior into their knowledge of how to anticipate hazards or if the effect was merely due to previewing the exact same driving simulator scenario that they would be tested on immediately after watching the video.

For the current study, we aimed to build upon previous research to identify which aspects of the RAPT and expert video condition (RAPT-EV) led to better hazard anticipation performance. To evaluate this, an additional group of participants were run using the same protocol as the previous study (Yamani et al., 2018). This new group of participants completed the RAPT protocol and watched a 10-minute video of the simulated driving scenario containing the anticipatory eye movements of a young driver who failed to correctly anticipate the hazards in any of the scenarios (RAPT-NV).

If the RAPT-NV group demonstrates similar improvements in hazard anticipation performance to the RAPT-EV group, then improvements in performance for both video conditions are due to participants previewing the driving scenario that they will be driving, not because they were able to successfully map and incorporate correct hazard anticipation glance behavior into their own mental models. However, if the RAPT-NV group's hazard anticipation performance is worse than RAPT-EV group, then modeling correct eye glance behavior leads participants to integrate the skills learned in RAPT more effectively.

METHOD

Participants

There was a total of 36 young drivers, recruited from the Old Dominion University undergraduate community (22 women, 14 men). All participants were between the ages of 18-21, an age range that would still have suboptimal hazard anticipation skills (Pradhan et al., 2005). Information for the participants' ages and years since licensure is displayed in Table 1. Twenty-four of these participants (from the RAPT-EV and RAPT only conditions) were part of a previous study (Yamani, et al., 2018). An additional 12 participants were recruited for the RAPT-NV condition. Participants held a valid US driver's license, had normal or corrected-to-normal vision and drove less than 10,000 miles since licensure. Participants were remunerated for their participation.

Table 1. Mean and standard deviation of age and years since licensure by group

	Age		Years since licensure	
	Mean	*SD*	*Mean*	*SD*
RAPT -NV	19.17	1.11	3.31	1.06
RAPT -EV	19.08	1.08	2.06	0.94
RAPT Only	19.58	1.24	3.59	1.26
Total	19.28	1.14	2.99	1.26

Apparatus

Driving simulator. A fixed-base STISIM Drive simulator was used for this study. The system included a Dell Studio XPS with Windows Vista x64 Enterprise, a gaming Playseat, and Logitech G27 racing wheel and pedals. To project the simulated environments, a DPL 1800 MP Front Projector paired with a 76-inch white smart board screen was used. Participants viewed the screen from a distance of about 177 cm. The driving simulator was paired with a surround speaker system that appropriately simulated environmental sounds, vehicle noise and doppler.

Eye tracker. An ASL Mobile Eye head-mounted eye tracker was used to track participants' eye movements throughout the drive. The system processed the recorded scene and eye images into a single video of the scene which displayed the location of gaze on each frame, represented by a crosshair. The eye position data were sampled at 60 Hz.

Scenarios. All participants completed a single driving simulator scenario that evaluated their hazard anticipation ability. The drive consisted of four separate hazard anticipation scenarios, each involving a potential latent hazard existing within a *target zone* (area of a roadway where a threat could materialize) as the driver navigates the *launch zone* (the section of the roadway where a driver should start scanning for latent hazards). No threat actually materialized in any of the scenarios. Each of the four hazard anticipation scenarios corresponded to a different driving environment (highway, town, rural, and residential). These scenarios were identical to those used in previous studies (Yamani et al., 2018; Yamani et al., 2016). An example of a scenario is shown in Figure 1.

Figure 1. Example of latent hazard anticipation scenario. The image on left shows a correct anticipatory glance to look for hazards that might be hidden by the bus. The image on right shows the glance behavior of a participant who failed to anticipate the latent hazard.

Training program. The RAPT-3 was used for all training conditions (RAPT only, RAPT-EV, and RAPT-NV). RAPT both assesses a driver's skill at being able to evaluate potentially unsafe situations and trains them to effectively visually scan areas of the driving scene that may contain latent hazards. RAPT-3 uses nine different scenarios that have been validated in previous studies, and has three sections –pre-test, training, and post-test. In the pre-test, drivers view a sequence of scenario snapshots from an on-road perspective and are asked to click on areas of the roadway where they would look – no feedback is provided. The training section presents a top-down view of the scenario with narrative explanations of the risk present in each scenario. Participants can then practice identifying the areas of risk on the scenario snapshots up to four times. The post-test section is identical to the pre-test; participants again view the nine scenario snapshots. The training program took approximately 40 minutes for participants to complete.

Eye movement video. In order to create the "expert" eye movement video, an experimenter drove through the scenario wearing the eye tracker. The crosshair from the eye tracker was superimposed onto the scene video. The experimenter displayed perfect anticipation (correctly glancing at the target zones while in the launch zones) in each hazard anticipation scenario. This video was shown to the RAPT-EV group. For the RAPT-NV group, an eye tracking video of a participant in Yamani et al. (2018) who failed to anticipate all the hazards (did not correctly glance at the target zones while in the launch zone) was used.

Procedure

This experiment used a one-way between-subjects design, with participants assigned to one of three training groups. The dependent variable was the average proportion of anticipatory glances made by each group. All participants provided explicit informed consent to participate in the study. Participants first completed the RAPT training. Following the completion of RAPT, two groups (RAPT-EV and RAPT-NV) watched a 10-minute video clip of the driving simulator scenarios they would complete in their evaluation drive. Participants in both the RAPT-EV and RAPT-NV groups were told that the video was of a drivers' eye movements and part of the training program, but they were not told whether the eye glances were that of an expert or a novice. The RAPT only group did not watch any additional video. Before starting the drive, participants were equipped with an eye tracker. After a 5-minute practice drive in the simulator, participants completed a single 10-minute evaluation drive, which consisted of four driving

environments (highway, town, rural, and residential), counterbalanced amongst participants. In each environment, there was one hazard anticipation scenario. Each experimental session lasted approximately 90 minutes.

The eye glances of the participants were recorded. Two independent (double blind) raters coded whether participants correctly looked at the target zone while driving through the launch zone, which was predetermined for each scenario. This protocol has been used in previous hazard anticipation evaluations (Unverricht et al., 2018) The parameters of each launch zone were discussed prior to coding. Videos that had been previously coded in Yamani et al. (2018) were re-coded by the raters to ensure consistency. Disagreements were discussed between the two raters to determine the proper coding.

RESULTS

Instead of using null hypothesis significance testing (NHST), we used default Bayesian analysis to compare differences between the groups. Unlike NHST, Bayesian analysis allows statistical evidence *for* and *against* an effect of interest. Bayes factor is a ratio of the likelihood that obtained data arose from a model that includes the effect of interest, to a model that excludes the effect (Rouder & Morey, 2012). We use B_{10} as a measure of evidence as suggested by Rouder & Morey (2012). According to guidelines put forth by Jeffreys (1961), Bayes factors between 1 and 3 provide only *anecdotal* evidence for an effect. Bayes factors between 3 and 10 provide *substantial* evidence for the presence of an effect.

The proportion of correct glances for each of the four hazard anticipation scenarios was averaged across training condition groups. The mean proportion of correct hazard anticipation responses is displayed in Figure 2.
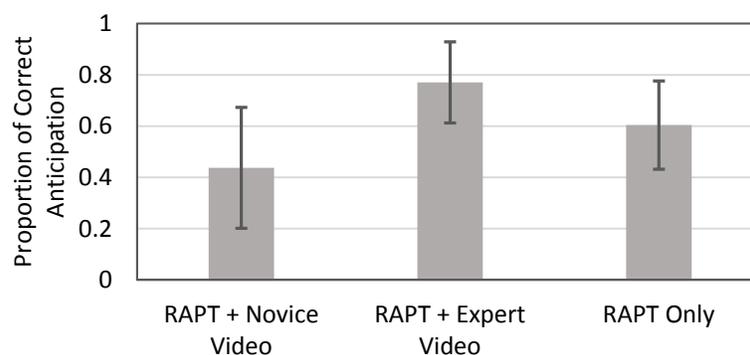
Figure 2. Mean proportion of correct anticipation and standard error by training group. Error bars represent 95% confidence intervals of training group mean.

A series of Bayesian t-tests were used to compare proportions of correct anticipation between the conditions. Interestingly, participants in RAPT-EV showed substantially better hazard anticipation than those in RAPT-NV, $M = .77$ vs. $.43$, $t(22) = 2.58$, $B_{10} = 3.51$. However, data indicate no substantial evidence for performance differences between RAPT-NV and RAPT-only, $M = .43$ vs. $.60$, $t(22) = 1.26$, $B_{10} = 1/1.52$, and between RAPT-EV and RAPT-only, $M = .77$ vs. $.60$, $t(22) = 1.57$, $B_{10} = 1/1.12$.

DISCUSSION

This experiment tested whether the near-ceiling hazard anticipation performance reported in Yamani et al. (2018) was due to the expert eye movements or an opportunity to preview the simulated driving that the expert eye movement video offered. Young drivers in RAPT-NV anticipated a lower proportion of latent hazards compared to their peers in the RAPT-EV condition. However, unlike the results of Yamani et al. (2018), after the videos were re-coded by new raters, there was no substantial evidence to suggest adding the expert eye glance video to the RAPT program improved hazard anticipation scores above and beyond those of RAPT-only condition. Nor was there any evidence indicating that RAPT-NV anticipated hazards any worse than the RAPT only group. This might reflect more stringent scoring criteria employed by the current coders than those coders in the original Yamani et al. (2018) study. Although this does not necessarily undermine the current study showing that latent hazard anticipation performance differs between the RAPT-EV and RAPT-NV, a replication study of Yamani et al. (2018) may help determine a more accurate estimate of the size of the expert eye movement video effect on latent hazard anticipation.

The difference between RAPT-EV and RAPT-NV suggest that merely previewing a video with eye movements of a driving scenario is not enough to increase hazard anticipation scores. However, the fact that the RAPT-NV group tended to perform worse than the RAPT only group raises a new set of questions. First, it is possible that young drivers are sensitive to interpreting eye glance behavior, and so they internalized the behavior that they watched regardless of the accuracy of the eye movements presented to them immediately before their evaluation drives. Alternatively, because the novice's tracked eye glances would have demonstrated behavior counter to what the young drivers had just learned in the RAPT module, drivers in the RAPT-NV group may have been confused by the conflicting training they were receiving. These possibilities could explain why watching a novice's eye glances created a detrimental effect.

One possible follow-up to this study could be to compare latent hazard anticipation between RAPT- EV and a condition of an identical video without the crosshair, allowing us to understand whether participants would perform better on the hazard anticipation task after previewing the scenario without glance behavior that conflicted with the skills they learned during the RAPT protocol. Another study could investigate the transfer effect of watching an expert's eye glances on a dynamic roadway that does not preview in any way the hazard anticipation scenarios that the trainees will be tested on afterwards. This could help explain if trainees who watch videos of proper scanning behavior are able to internalize and generalize the skills to novel situations. If young drivers can successfully internalize expert eye glances into their mental model of how they approach latent hazards, this would provide support for incorporating expert eye movements into existing or novel hazard anticipation training programs. Following additional confirmatory evidence, there may be some merit to incorporating some form of eye movement-based training in general drivers' education. Further, there may be broader implications for targeted skill-specific training in other domains such as manufacturing, healthcare and aviation. From the perspective of an automotive designer, a better understanding of how eye movements can be internalized may offer considerations for better interface interaction design and could offer pathways for collaborative user environments (e.g., Yamani, Neider, Kramer, & McCarley, 2017).

There are several limitations to this study. The instability of the anticipatory glance ratings from one set of coders to another may mean that in future studies it would be best to use more than two coders. Or alternatively, it may be beneficial to use automated Areas of Interest coding to improve the accuracy and reliability of the ratings. The current study was conducted on a driving simulator and therefore an on-road study would be valuable at providing external validity. There may be extended benefits of such training for other population demographics such as teen drivers and perhaps aging drivers. More research is needed to substantiate this further.

REFERENCES

Jeffreys, H. (1961). *Theory of probability* (3rd Ed.). Oxford, UK: Oxford University Press.

Insurance Institute for Highway Safety. (2017). *Fatality Facts: Teenagers*. Retrieved from https://www.iihs.org/iihs/topics/t/teenagers/fatalityfacts/teenagers

McDonald, C. C., Goodwin, A. H., Pradhan, A. K., Romoser, M. R. E., & Williams, A. F. (2015). A Review of Hazard Anticipation Training Programs for Young Drivers. *Journal of Adolescent Health*, *57*(1), S15–S23. https://doi.org/10.1016/j.jadohealth.2015.02.013

McKnight, A. J., & McKnight, A. S. (2003). Young novice drivers: careless or clueless? *Accident Analysis & Prevention*, *35*(6), 921–925. https://doi.org/10.1016/S0001-4575(02)00100-8

NHTSA. (2018). *Traffic Safety Facts, Young Drivers* (No. DOT HS 812 498) (p. 8).

Pradhan, A. K., Hammel, K. R., DeRamus, R., Pollatsek, A., Noyce, D. A., & Fisher, D. L. (2005). Using Eye Movements to Evaluate Effects of Driver Age on Risk Perception in a Driving Simulator. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(4), 840–852. https://doi.org/10.1518/001872005775570961

Pradhan, A. K., Pollatsek, A., Knodler, M., & Fisher, D. L. (2009). Can younger drivers be trained to scan for information that will reduce their risk in roadway traffic scenarios that are hard to identify as hazardous? *Ergonomics*, *52*(6), 657–673. https://doi.org/10.1080/00140130802550232

Rouder, J. N., & Morey, R. D. (2012). Default Bayes Factors for Model Selection in Regression. *Multivariate Behavioral Research*, *47*(6), 877–903. https://doi.org/10.1080/00273171.2012.734737

Taylor, T., Masserang, K., Pradhan, A., Divekar, G., Samuel, S., Muttart, J., … Fisher, D. (2011). Long-Term Effects of Hazard Anticipation Training on Novice Drivers Measured on the Open Road. In *Proceedings of the 6th International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design:*(pp. 187–194). Olympic Valley-Lake Tahoe, California, USA: University of Iowa. https://doi.org/10.17077/drivingassessment.1396

Unverricht, J., Samuel, S., & Yamani, Y. (2018). Latent Hazard Anticipation in Young Drivers: Review and Meta-Analysis of Training Studies. *Transportation Research Record*, 0361198118768530.

Yamani, Y., Bıçaksız, P., Palmer, D., Hatfield, N., & Samuel, S. (2018). Evaluation of the Effectiveness of a Gaze-Based Training Intervention on Latent Hazard Anticipation Skills for Young Drivers: A Driving Simulator Study. *Safety*, *4*(2), 18. https://doi.org/10.3390/safety4020018

Yamani, Y., Neider, M. B., Kramer, A. F., & McCarley, J. S. (2017). Characterizing the efficiency of collaborative visual search with systems factorial technology. *Archives of Scientific Psychology*,*5*(1) 1

Yamani, Y., Samuel, S., Knodler, M. A., & Fisher, D. L. (2016). Evaluation of the effectiveness of a multi-skill program for training younger drivers on higher cognitive skills. *Applied Ergonomics*, *52*, 135–141. https://doi.org/10.1016/j.apergo.2015.07.005