

Can a driving simulator assess the effectiveness of Hazard Perception training in young novice drivers?

T. Oron-Gilad Y. Parmet

*Department of Industrial Engineering and Management – Ben-Gurion University of the Negev,
P.O. Box 653, Beer-Sheva, Israel
email: orontal@bgu.ac.il, iparmet@bgu.ac.il*

Abstract

Hazard awareness skills of young-novice drivers are receptive to training. Yet, there is no consensus on an optimal training program or acceptable measures to assess intervention effectiveness. The current study aimed to develop a driving Simulator based Hazard Perception Test (S-HPT) for assessing changes in awareness of hazards among trained young-novice driver, relative to a control group, and a group of experienced drivers (gold standard). Two urban environments were chosen for testing; residential district with narrow roads, and business district with wider roads. Fourteen scenarios consisting of potential situations known to differentiate between untrained young-novice and experienced drivers (e.g., limited field of view) with and without materialized situations were created. Data from 39 young-novice and 6 experienced drivers was analyzed. Using an allocation algorithm each young participant was assigned to one of three training interventions (AAHPT-Active, Hybrid or RAPT) or control. Driving speed was chosen as the measure of analysis since it indicates whether drivers have spotted a hazard, and what they chose to do to avoid it. It was sampled every 3 meters, approximately 400 times per scenario per driver group. A respective curve was fitted for each group. Differences in speed among driver groups were examined and comparisons between potentially hazardous scenarios and materialized ones were made. Results revealed that scenarios that required braking but not to a complete stop were the most diagnostic. The use of two environments demonstrated differences in group behaviours that would have not been seen if utilizing only one environment. Use of two comparisons, to control and to experienced drivers was reassuring as the results complemented. Where the resemblance of trainees was higher to control, they tended to resemble the experienced group less. The Active trainees bore more resemblance to the control in the residential scenarios while the RAPT and Hybrid bore more resemblance to the control in the business district scenarios. To conclude, generation of a variety of scenes in multiple traffic environments with- and without materialized hazards is a promising way for hazard awareness simulator test development. Group-related metrics and analyses are essential to investigate which conditions discriminate among driver groups.

Keywords – hazard awareness, training, driving simulator, young-novice drivers, AAHPT, RAPT

1. Introduction

A large body of evidence supports the notion that the ability of young drivers to identify hazardous situations is inferior to that of experienced drivers (e.g., [1, 2, 4, 10, 13, 15, 20, 22-24]). Difficulties to identify potential hazards among young-inexperienced drivers can be attributed to the fact that they are incapable of integrating elements within the traffic environment to create a holistic picture of the situation at hand [1].

It is agreed among the scientific community that skills related to hazard perception can be improved by training but little consistency exists among training interventions or in the way they are evaluated. McKenna and Crick [14], for example, developed a training methodology that included short road films, while instructing participants to maintain a forward look towards the driving direction. On different time points, the film was paused, and participants were asked to predict the possible future scenarios. A different approach for practice was suggested by Mills et al. [18] who compared performance of participants who received instructions in a classroom combined with short road films, and participants who drove real-world situations and were instructed by an authorized driving instructor. Both techniques reduced latencies in a video-based hazard perception test, though there has been a greater reduction in latencies for drivers who underwent both techniques. Error training is another approach, Wang et al. [26] examined simulation-based error training and video based guided error training. Measured in a simulator driving test, simulation-based error training seemed more helpful for novice drivers in anticipating and detecting hazards. Commentary is another technique requiring participants to provide verbal commentary while driving, creating a verbal protocol which combines what the driver is seeing and thinking, with what they are planning to do to avoid potential hazards. Crundall et al. [7] compared simulated driving behavior of one group of drivers who were trained in commentary driving to a control group. Results showed that the trained group had fewer crashes, reduced their speed sooner on approach to hazards, and applied pressure to the brakes sooner than untrained drivers.

Encouraging evidence regarding improved hazard perception abilities measured in laboratory settings were reported by two independent training methodologies, RAPT (Risk Awareness and Perception Training: Fisher et al. [10, 20]) and AAHPT (Act and Anticipate Hazard Perception Training: Meir et al. [16, 17]). The AAHPT envisions that exposing young-inexperienced drivers to a vast array of video clips of real-world traffic situations can facilitate their skills of anticipating upcoming hazardous situation and such knowledge can be transferred to similar but novel situations. RAPT assumes that teaching drivers where to search for hazards, according to some principles, can enhance their anticipation abilities in other, novel situations. In the evaluation phase, AAHPT studies used a video-based evaluation test [17] while RAPT evaluations utilized a driving simulator test. This makes it difficult to compare their effectiveness on similar circumstances. Yet, despite their inherently different training and testing methodologies there was evidence that some of the findings of both studies shared resemblance. For example, Chan et al. [6] found that participants who underwent RAPT demonstrated sensitivity to the dangers of pedestrians, intersections, and limited field of view. Meir et al. [17] reported that drives who underwent the AAHPT-Hybrid training demonstrated increased sensitivity to pedestrians and limited field of view, both in terms of responses to hazards and eye scanning patterns. These findings suggest that both training methodologies increase the sensitivity of drivers to specific materials that were learned during training but it is not clear whether there is a negative trade-off with paying attention to other situations that were not emphasized during training.

Thus, it is evident that in light of the many interventions that have been available, and in light of the variation in evaluation tests, there is a pursuing need to develop more consistent evaluation techniques and tests. It remains unclear which of the training paradigms may be superior or which type of test measurements may best estimate skill improvement [24]. Meanwhile, there is a growing interest in using driving simulators for assessment, particularly because they can record complex behavioral responses continuously. Underwood et al. [23] argued that driving related measures derived from the simulator such as speed and braking provide a behavioral signature that indicates that a driver has spotted the hazard, and also what behavior he has chosen to avoid it.

Such behavioral signatures distinguish between experienced and novice drivers, as well as between groups of learner drivers (e.g., [7]). Fisher et al. [9] found that young drivers, as opposed to experienced drivers, have difficulty to adjust their driving speed as a response to traffic signs and changes in the road. For example, young drivers tended to use the brake of the vehicle and did not reduce their speed gradually as opposed to experienced drivers, in response to a “Stop” sign on a road curve. Wang et al. [25] have asked eight experienced and eight young-inexperienced drivers to drive five simulated hazardous scenarios in a driving simulator. They found that experienced drivers drove slower than young-inexperienced drivers when they approached a potential hazard and that their reaction to imminent threats was more cautious than that of young-inexperienced drivers. Oron-Gilad et al. [19] used group based assessment metrics of driving speed to analyze driving speed performance differences between untrained young-novice drivers and experienced drivers in the driving simulator. They reported that experienced drivers were more anticipating of hazardous events particularly in business district urban area with many uncontrolled intersections. Experienced drivers approached the uncontrolled intersections while reducing their speed, and they tended to better anticipate the behavior of a merging lead vehicle. Furthermore, they reported that although all drivers attended to materialized events (e.g., a pedestrian on the road or a merging vehicle), experienced drivers as a group, attended to these events earlier in the trajectory. Regan et al. [21] examined whether 18 years-old trained learner drivers drive better than a similar group of untrained learner drivers. Participants drove 7 simulated scenarios on one week and then again four weeks after training. The authors reported that in 50% of the traffic scenarios trained drivers performed better and tended to brake earlier in response to a hazardous event.

In sum, young-novice drivers are less aware of hidden and potential hazards than experienced drivers (e.g., [2, 10]). For this reason, it has been hypothesized that young-novice drivers will manage their speed in a less consist way than experienced drivers when encountering a hazardous situation (e.g., [9, 19]). The literature provides sufficient evidence that untrained young-novice drivers are less aware of potential hazards (e.g., [7]) than trained ones. The driving simulator allows to examine whether increased awareness to potential or materialized hazardous situations occurs, and if so, how it is reflected in driving behaviour. If the training intervention is successful, one should expect trained drivers to be more aware of potential hazards and more prepared toward materialized hazard, i.e., resemble the experienced driver group more. Contrarily, if the training intervention is less effective, no particular difference between young-novice trainees and control should be apparent. The development of the S-HPT was aimed toward creating a diagnostic way to identify how training interventions affect driving speed behavior in light of the trajectory and the driving scenes.

2. Development of the driving Simulator based Hazard Perception Test (S-HPT)

The S-HPT was developed to explore whether training interventions increase awareness toward materialized and/or potential hazardous situations among young-novice drivers who underwent hazard awareness related training. It was of particular interest to examine which types of scenes are diagnostic and separate among the groups. Then, these differences in driving behaviour can be further analyzed in light of the content and focus of the intervention.

At first, scenarios that consisted primarily of potential hazards, shown to distinguish young-novice drivers' behaviour from experienced drivers (e.g., [3, 19]) were designed in two urban environments; a residential area consisting of narrow single lane road populated with pedestrians and parked vehicles, and a business district less populated environment, one or two lanes in each direction, often with a separation island, and more vehicles moving at relatively higher speed.

Use of multiple traffic environments is common in hazard perception tests (e.g., [17]) and shown to be meaningful when examining differences between novice and experienced drivers in simulators [8]. As demonstrated in Oron-Gilad et al. [19], the business district area provides more degrees of freedom in choosing driving speed, yielding larger differences in group homogeneity of speed between untrained young-novice drivers and experienced drivers (the group homogeneity of speed analysis examines the variation among driver groups for each specified interval along the scenario, see Yona et al. [27] for a detailed description). In the current study, in order to evaluate the change in sensitivity of young-novice drivers to potential events, sets of similar scenarios with- and without materialized events were chosen and compared.

Secondly, the method of analysis is important. Similar to Oron-Gilad et al. [19], the focus of the analysis was on examining driving speed over the course of an entire driving scenario, along the trajectory of the entire scene. In contrary to focusing on particular planned events (e.g., [7]) and in differ from an event-based analysis method that explores only segments in the scenario that the researcher has a-priori identified as meaningful, the use of a curve fitting technique allows to broadly examine changes in driving speed of a group of drivers over the course of the entire trajectory. This offers a solution to the problem that there are always more potential events in a scenario than the ones initially planned by the researchers [2] which makes it difficult to separate “critical-hazard anticipation” areas in a scenario from “less critical” ones [24]. Furthermore, this type of examination minimizes the impact of disparity between simulator and real-road driving since it takes into account that driving a simulator has fixed consistent effects on driver behaviour across the entire drive. These fixed effects only change the start and the end points of the measurement, but not the trends of the lines. Since trained and untrained drivers were included in the analysis, emphasizing the need for a sensitive measurement mechanism, as performance differences among groups of young-novice trainees and control may be less prominent than the ones shown previously between untrained young-novice drivers and experienced drivers (e.g., [19]). It was therefore decided to use the curve fitting technique of cubic smoothing spline [5, 11]. As opposed to the piecewise linear regression model used previously, the cubic smoothing model does not require preliminary assessments of the number of breakpoints in the model, and also one can, if wanted modify the degree of smoothness, although in this case it was done entirely by the algorithm without manual intervention. It was hypothesized that the business district area which provides more flexibility in choosing the driving speed will reveal larger differences among driver groups, while the residential area that is more restricting in nature, will provide smaller differences and be of less diagnostic value. Also, it was hypothesized that as trainees resemble the control group more, their hazard perception training intervention is less effective and oppositely, as trainees resemble the experienced drivers more, and their training is more effective. If no such differences emerge, it is indicative that the particular scenario is not sufficiently diagnostic and changes should be made to future versions of the S-HPT.

3. Experimental scenarios

The use of a variety of traffic environments has been one of the principles of evaluation applied in previous studies [17], as the driving environment dictates the type and frequency of occurrence of various hazardous situations. Therefore, the drive consisted of a fixed set of 14 urban scenarios, 8 business area scenarios and 6 residential area scenarios that were merged into a single drive (18 km long) while incorporating transition segments between scenarios. Each scenario was approximately 1,220 m long. The order of presentation varied among participants. Half of the participants in each group drove the business areas scenarios first and the other half drove the

residential scenarios first. Within each area the order of scenarios changed randomly among participants. No speed signs were shown in the simulation so that drivers will be less attuned to thinking about how they are expected to drive, and manage their speed more according to the travel speed that the traffic environment dictates. Drivers were informed of the legal speed limit in both areas (50 km/h). Other traffic signs were present in the simulation. Rear window mirror was present at all scenes.

- *Residential scenarios.* The aim was to examine driver performance in a populated environments consisting of narrow (lane width: 3.65 m) single lane roads where pedestrians are present and parked vehicles are obscuring part of the visual field of view and restricting driving-maneuver possibilities. Two pairs of scenarios are detailed in Table 1 and sample snapshots are shown in Figure 1
- *Business district scenarios.* The aim was to examine driver performance in main commercial area urban roads, with one or two lanes in each direction and more traffic (i.e., dynamic vehicle elements) moving at a relatively higher speed than on residential roads. Two pairs of scenarios are detailed in Table 1 and snapshots are shown in Figure 2.
- *Training scenario.* A separate acclimation 5 minutes long scenario, aimed to allow drivers to adjust to the vehicle's controls and display, consisted of driving in both business district and residential areas but without any materialized hazards.

Tab. 1 - Description of scenarios and events

Scenarios	Description	Event
R1-R2	Residential road with parked vehicles either on the right side of the road or on the left, but not simultaneously on both sides.	In R1 one of the parked vehicles pulls out into the driver's lane (without signalling) which requires immediate braking (materialized). In R2, using the same road, there were no planned events that required an immediate brake (potential)
R3-R4	Residential road, the driver approaches a midblock crosswalk.	In R3 the crosswalk is partially obscured by parked vehicles. It requires slowing down. Immediately following the crosswalk, a parked vehicle pulls out into the driver's lane which like in R1, requires immediate braking (materialized). In R4 the crosswalk is entirely visible to the driver (potential).
U1-U2	An urban main road with a sharp curve.	In U1 a stopped vehicle is located behind the apex of the curve and is obscured from the driver until passing the apex. Looking across the curve is not possible due to vegetation (materialized). In U2 there were no other cars in the scene.
U3-U4	An urban main road with a midblock separation. The driver follows a bus. The bus stops at the bus station.	In U3 a pedestrian from the curb on the opposing side runs toward the bus station (materialized). In U4 no planned events happen (potential).



Fig. 1 - Sample snapshots of events in residential scenarios. Top-left: a parked vehicle (marked with an arrow for emphasis) pulls out into the driver's lane (R1). Top-right: a residential road (R2). Bottom-left: a crosswalk partially obscured by parked vehicles (R3). Bottom-right: a clear view of a crosswalk (R4)

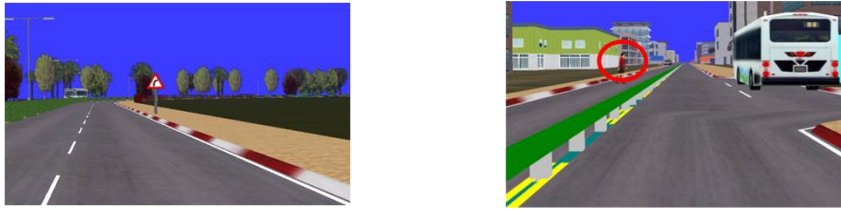


Fig. 2 - Sample snapshots of events in urban scenarios. Left: a curve in the road (U1-U2) in U1 there was a parked vehicle beyond the curve which could not be seen due to the vegetation. Right: a bus parked in the station and a pedestrian (marked by a red ellipse) crossing the road catch the bus (U3) in U4 there was no pedestrian

4. Method

4.1. Participants

Thirty nine young- novice drivers, 17-18 year-olds with less than three months of driving experience, still within their obligatory escort period (must drive accompanied by an adult driver with more than 7 years of driving experience). Six experienced drivers (with more than 8 years of driving experience, mean age 26) were used as a gold standard for comparison and analysis of the testing phase.

Each young participant filled a demographic and the sensation seeking questionnaire [28]. An allocation algorithm to the experimental groups (training or control) was then used to create matching, homogenous groups of young-novice drivers, to ensure to a high extent that the groups differ only by the specific treatment they have each received. The allocation process took into account key demographic variables found to be confounding with driver experience (e.g., [11]) according to descending priority: (1) sex, (2) level of driving exposure (self-report of the average number of hours the participant drove each week), and (3) the SSS score.

4.2. Training interventions and control

- *AAHPT Active.* Participants undergoing this condition were asked to engage in a hazard detection task. That is, they were required to observe the 63 HP training movies, randomly presented, and actively respond by pressing a response button each time they detected a hazard. Participants were instructed to regard the button press as a representation of any action (e.g., braking, swerving the wheel, honking the vehicle horn) that they would have taken in order to prevent a crash in the given situation. At the end of each movie, participants were instructed to fill in blank fields (which were equal to their number of responses in the movie) with a description of the hazard instigator that triggered their response. This procedure was aimed to resolve ambiguities regarding the reason for initiating each response. Notably, although the entries were open-ended, most participants in almost all movies did not indict any recall issues; almost all drivers tended to recall the hazard instigators (and rarely to mix the chronological order). In these cases we could use the eye-movements data to decipher their intended response. See Meir et al. [17] for details.
- *AAHPT Hybrid.* This training mode was a combination of the AAHPT Instructional component, where the trainee is exposed to various definitions for hazards followed by examples of traffic-scene movies, and the AAHPT Active intervention. It was constructed from two parts: (1) Part I contained a background overview utilizing 23 of the HP training movies to convey knowledge concerning HP and Part II was similar to that of the Active training and

contained a hazard detection task participants were required to employ on the additional 40 HP training movies. Notably, this design preserved the total training duration. See Meir et al. [17] for more details.

- **RAPT.** The Israeli RAPT (I-RAPT) is based on the RAPT [10, 20] with several modifications. As in the original RAPT, pre- and post- training evaluations were conducted, in which a sequence of snapshots from each traffic scenario was presented to participants. Participants were then instructed to imagine themselves as drivers. Using the mouse they were then asked to click on regions in the view where they would look for information which could reduce their likelihood of a crash in case they were actually driving the scenario.
- **Control.** The control group was presented with a tutorial concerning road safety, unrelated to hazard awareness.

4.3. The STI-SIM driving simulator

The STI-SIM driving simulator (Systems Technology, Inc.) integrated into a full-size Cadillac CTS sedan provided the driver with a look and feel of driving a real car. The visual display of the road was projected on a 7 m diameter round screen at a distance of 3 m from the driver's eyes, providing the driver with a true horizontal field of view of ~150 degrees on a scale of 1:0.8. The experimenter sat in a separate control room where computer screens allowed monitoring the driver and the scene.

4.4. Procedure

Each young-novice participant participated in two sessions one week apart: (1) training or control, (2) testing in the driving simulator. Based on the allocation algorithm aforementioned, each young-novice participant was designated to one of four experimental conditions: AAHPT-Active, AAHPT-Hybrid, RAPT or control. In the test session, after debriefing, drivers had a short adaptation period to the simulator in order to get familiar with the driving simulator's steering wheel and pedals. Then, they were asked to drive through the entire sequence of 14 traffic-scenarios as if they were driving in real world situations while maintaining the legal speed (50 km/h) and obeying the traffic rules. Upon completion of the drive, the experiment ended and participants were debriefed and compensated for their time.

5. Results

5.1. Driving speed curves

The focus in this article is on driver-group behavior throughout a scenario. Therefore we did not take one measure for each driver per each scenario, but instead we followed driving speed throughout the scenario, for each participant and each scenario. The simulator sampled each driver's longitudinal velocity every six feet (a time interval of about half a second when driving at legal speed of 50 km/h). This sampling was used for the analyses. The average speed among individuals of the same group of drivers (i.e., control/ AAHPT Active/ AAHPT Hybrid/RAPT/experienced) was then calculated for each particular interval. As a result of this stage five series of data points were obtained for each sampling point in each scenario. Since scenarios were about 1200 m long, approximately 400 sampling points were obtained in each scenario for each group. The last step of the analysis was to fit cubic smoothing spline, fitting a smooth curve to a set of noisy observations [5, 11] to each group, shown as the solid lines in Figures 3-5.

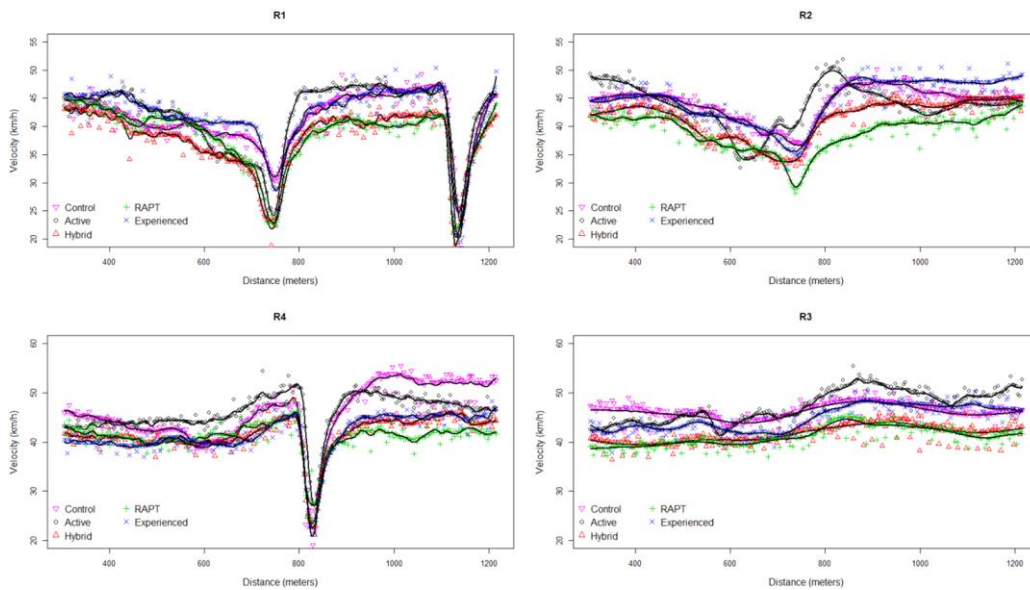


Fig. 3 - Plots of longitudinal velocity in the residential scenarios by distance and driver groups. The solid lines represent the fitted curve, dots represent each driver. Marked areas are enlarged in Figure 4

A statistical test was then conducted to examine whether the five separate curves, fitted for each group, could be replaced by a single curve (i.e., that all groups chose their speed in the same way). This test showed that for all 8 scenarios, the separate group curves could not be combined into one: R1:F (284, 1165) = 32.4, $p < .00001$, R2:F (149, 1331) = 101.2, $p < .00001$, R3:F (264, 1189) = 40.9, $p < .00001$, R4:F (103, 1387) = 82.8, $p < .00001$; U1:F (200, 1279) = 107.0, $p < .00001$, U2:F (103, 1383) = 124.4, $p < .00001$, U3:F (204, 1255) = 34.9, $p < .00001$, U4:F (101, 1382) = 111.1, $p < .00001$. Since the groups were behaving differently, additional descriptive examinations considering differences in pattern among groups were conducted, and detailed in the following sections.

5.2. Residential scenarios

Figure 3 shows the distribution of longitudinal velocity sampling points per each group, per sampling point along scenarios R1-R4. The solid lines are the fitted curves for the longitudinal velocity aforementioned. Three types of examinations were made: 1) comparison of potential event drives (R2 and R4) with their corresponding materialized event drivers (R1 and R3), respectively; 2) comparison among the various training groups and the control; and 3) comparison among the various training groups and the experienced.

- *Response to potential vs. materialized events.* With regard to the differences between scenarios that consisted of a materialised event versus those consisted only of potential hazards, it is evident that all drivers responded to the materialized events (R1 and R3). Nevertheless, in scenario R2, it can be seen that although in the planning of the scenario no events the researchers intended this scenario to be smooth and require no braking, potential events can also cause a reduction in speed (due to the road conditions or constraints) that they were not anticipated in advance. On the contrary, in scenario R4 there were changes in speed along the drive, as anticipated. It is therefore interesting to examine whether all driver groups respond to

the braking (whether due to materialized events as in R1 and R3 or potential one as in R2) in a similar way when they resume their speed following the slow down. Figure 4 shows zooms into those specific areas where differences among driver groups' patterns are apparent, but not due to the type of hazard (R1 vs. R2 show a similar trend).

- *Comparison between trained driver groups and control.* From Figure 3 it is evident that the control group tended to drive faster than the RAPT and Hybrid training throughout all scenarios but not necessarily different than the Active trainees who bared more resemblance to the control. Figure 4 further emphasizes that the Active group drove in a more risky way than the control in R1 and R2, i.e., they resumed their speed earlier in the drive than the other groups. Contrary, in R4, following the materialized event, the control group tended to speed more than the Active but there was still a resemblance between the two groups.
- *Comparison between trained driver groups and experienced.* From Figure 3 can be seen that the Active group tended to drive faster than the experienced drivers while the RAPT and Hybrid groups tended to drive consistently slower. Figure 4 further emphasized that the Active groups bore less resemblance to the experienced following a braking incident. RAPT members seemed to be over cautious particularly in R2.

5.3. Business district scenarios

Figure 5 shows the distribution of longitudinal velocity sampling points per each group, per sampling point along scenarios U1-U4. The solid lines are the fitted longitudinal velocity aforementioned. Three types of examinations were made: 1) comparison of potential event drives (U2 and U4) with their corresponding materialized event drivers (U1 and U3), respectively; 2) comparison among the various training groups and the control; and 3) comparison among the various training groups and the experienced.

- *Response to potential vs. materialized events.* It is evident that all drivers responded to the U3 materialized event with an immediate brake. In U1, the curve and the parked vehicle generated a less pronounced response, and more dispersion among groups. Scenarios U2 and U4 consisting only of potential hazards, somewhat resembled the patterns seen in U1.
- *Comparison between trained driver groups and control.* From Figure 5 it is evident that the control group tended to drive closer to the RAPT and Hybrid training throughout scenarios U1 and U2 and U3. The Active group tended to drive much faster than the control on these scenarios (in U1 sometimes even more than 20 km/h faster). In U4 the control drove slightly faster than all the other groups, and also accelerated more after the materialized event, though they were not necessarily different than the Active trainees.

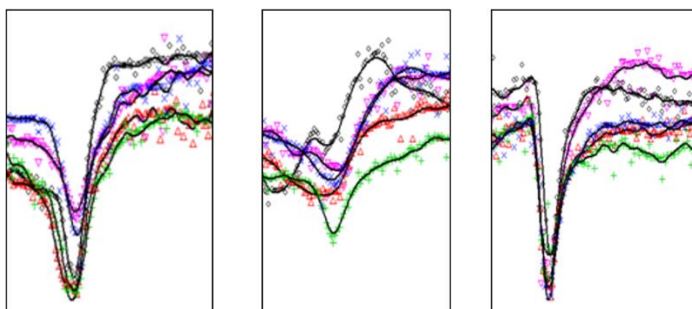


Fig. 4 - From left to right, zoom into particular areas in plots R1, R2 and R3, respectively

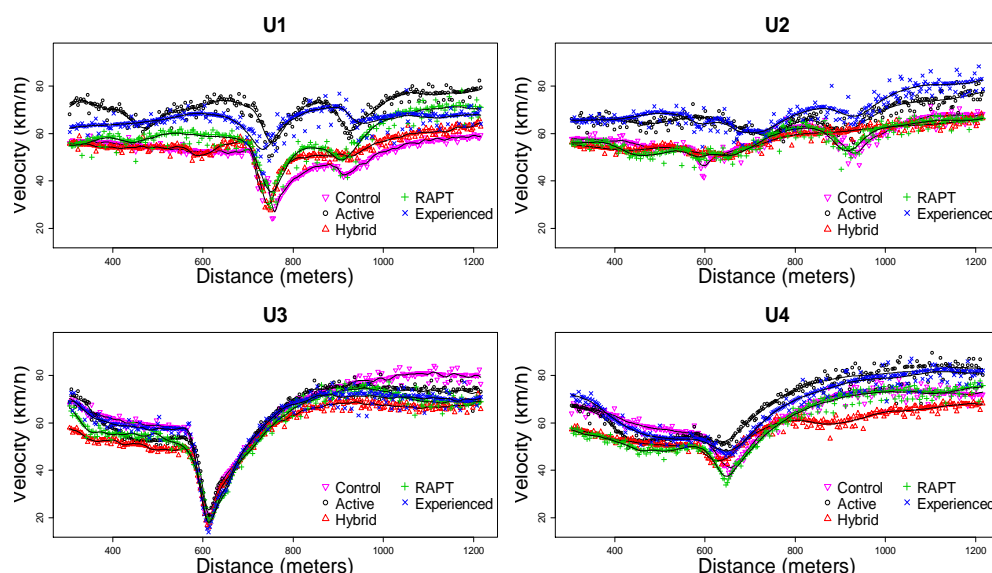


Fig. 5 - Plots of longitudinal velocity in the urban scenarios by distance and driver groups. The solid lines represent the fitted curve, dots represent each driver

- *Comparison between trained driver groups and experienced.* From Figure 5 it can be seen that the Active group tended to drive in a higher resemblance to the experienced drivers than the Hybrid and RAPT trainees who drove slower. In U1 and U2 one can almost see a trend for two separate clusters (experienced and Active vs. the other groups).

6. Discussion

The S-HPT was developed in order to explore whether training can increase awareness to potential hazardous situations among trained young-novice drivers. Examination of the way drivers continuously manage their speed in the simulator is one alternative to evaluate the simulator test's sensitivity and to select the more diagnostic scenarios. The results indicate that, as noted also by others [8, 17] it is indeed important to examine more than one traffic environment, as well as, to examine both potential hazards and materialized events. These issues are discussed in more detail.

It was hypothesized that the business district area which provides more degrees of freedom in choosing the driving speed will reveal larger differences among driver groups, while the residential area that is more restricting in nature, will provide smaller differences. This was indeed the case in Oron-Gilad et al. [19] where more disparity in driving speed between young-untrained and experienced drivers was found in the business district scenarios. However, the results of both residential and business district scenarios were informative and synergetic, as the residential scenarios added information that could not have been available from the business district ones. From the business district scenarios one could have concluded that the Active trainees bore the most resemblance to the experienced drivers' gold standard. Yet, an examination of the residential scenarios revealed a different pattern of less resemblance to the experienced group, implying that the Active trainees were perhaps over confident and more risk-prone in proximity to events (or potential events) that required braking.

Complementary to that, the active trainees bore more resemblance to the control in the residential scenarios while the RAPT and Hybrid bore more resemblance to the control in the business district scenarios. Finally, with regard to the development of the scenarios, oftentimes, participants identify more events than what was initially defined by the developers of the test (e.g., [2, 24]), this was demonstrated in R2, where participants braked although there was no apparent form of materialized hazard such as an approaching vehicle.

Furthermore, even though all drivers responded to materialized events, it became evident that there were difference among the groups in the way they recovered from a braking event and this type of recovery was oftentimes informative (e.g., the first event in R1, R2 and U1). The cases that seemed less informative were the ones that required a complete stop of the vehicle (e.g., the second event in R1 and U3) where no variation among groups could be found. Also, scenarios with potentially hazardous scenes were not all the same, R2 for example did require some slowing down and use of brakes while R3 and U4 did not. The latter two can therefore be used to provide a more general examination of speed handling across a distance of a particular environment.

All in all, the set of eight scenarios provided here showed initial evidence that the generation of a variety of scenes in each traffic environment with- and without materialized hazards is a promising way to go. Use of group-related metrics such as curve fitting of driving speed per group, are essential to investigate which conditions discriminate among driver groups with regard to general patterns of driving behaviour such as speed. Use of two types of comparisons, to control and to experienced drivers was also reassuring as the results complement, and where the resemblance of trainees is higher to control, they tend to resemble the experienced group less. Finally, there were pronounced differences among the trainee groups, findings which require delving back into the content of the training programs and inquiring the sources of certain behaviours. This type of investigation is beyond the scope of this manuscript.

7. Conclusions

S-HPTs are a promising way to evaluate the effectiveness of hazard awareness training interventions. However, the development of a simulator based test is not trivial and scenarios must be chosen carefully. It is recommended to use more than one traffic environment and provide similar traffic conditions with- and without- materialized hazards. Events that require a complete stop seem to be less diagnostic than events that require an immediate brake but not a complete halt. Driving speed as the simulator derived measure can provide diagnostic outcomes. Use of metrics that allow comparison among groups (such as the curve fitting method) are essential for the assessment of the scenario's diagnostic value.

Acknowledgements

This study was supported by the Israeli National Road Safety Authority.

References

1. Benda, H.V., Hoyos, C.G., (1983). Estimating hazards in traffic situations. *Accident Anal Prev*, 15, 1-9.
2. Borowsky, A., Oron-Gilad, T., Shinar, D. (2010). Age, skill and hazard perception in driving. *Accident Anal Prev*, 42, 1240-1249.
3. Borowsky, A., Oron-Gilad, T. (2013). Exploring the Effects of Driving Experience on Hazard Awareness and Risk Perception via Real-Time Hazard Identification, Hazard Classification, and Rating Tasks. *Accident Anal Prev*, 59, 548-565.
4. Brown, I.D., Groeger, J.A. (1988). Risk perception and decision taking during the transition between novice and experienced driver status. *Ergonomics*, 31, 587-597.

5. Chambers, J.M., Hastie, T.J. (1992). *Statistical Models in S*, Wadsworth, Brooks/Cole.
6. Chan, E., Pradhan, A.K., Pollatsek, A., Knodler, M.A., Fisher, D.L. (2010). Are driving simulators effective tools for evaluating novice drivers' hazard anticipation, speed management, and attention maintenance skills?. *Transport Res F-Traf*, 13(5), 343-353.
7. Crundall, D., Andrews, B., van Loon E., Chapman, P. (2010). Commentary training improves responsiveness to hazards in a driving simulator. *Accident Anal Prev*, 42, 2117-2124.
8. Dorn, L., Stannard, J. (2006). Simulator performance differences between experienced and novice bus drivers. *Adv Transport Stud, Special Issue*.
9. Fisher, D.L., Laurie, N.E., Glaser, R., Connerney, K., Pollatsek A., Duffy S.A. Brock J. (2002). Use of fixed-base driving simulator to evaluate the effects of experience and PC-based risk awareness training on drivers' decision. *Hum Factors*, 44(2), 287-302.
10. Fisher, D.L., Pollatsek, A.P., Pradhan, A. (2006). Can novice drivers be trained to scan for information that will reduce their likelihood of a crash? *Inj Prev*, 12, 25–29.
11. Green, P.J., Silverman, B.W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. Chapman and Hall.
12. Heino, A., van der Molen, H.H., Wilde G.J.S. (1996). Differences in risk experience between sensation avoiders and sensation seekers. *Pers Indiv Differ*, 20, 71–79.
13. Horswill, M.S., McKenna, F.P. (2004). Drivers' hazard perception ability: Situation awareness on the road. In S. Banbury and S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 155–175). Aldershot, UK: Ashgate.
14. McKenna, F.P., Crick, J.L. (1994). Hazard perception in drivers: a methodology for testing and training. Department of Transport, TRL CR313, Transport and Road Research Laboratory, Crowthorne, Berkshire.
15. McKenna, F.P., Horswill, M.S., Alexander, J. (2006). Does anticipation training affect drivers' risk taking?. *J Exp Psychol-Appl*, 12(1), 1-10.
16. Meir, A., Borowsky, A., Oron-Gilad, T., Parmet, Y., Shinar, D. (2010). Act and anticipate hazard perception training for young-inexperienced drivers, In P. Vink and J. Kantola (Eds.),
17. Meir, A., Borowsky, A., Oron-Gilad T. (2013). Formation and Evaluation of Act and Anticipate Hazard Perception Training (AAHPT) intervention for young-novice drivers. *Traffic Inj Prev*, 15(2), 172-180.
18. Mills, K.L., Hall, R.D., McDonald, M., Rolls, G.W.P. (1998). The effects of hazard perception training on the development of novice driver skills. Report to Department Environment, Transport & Regions.
19. Oron-Gilad T, Yona, O., Borowsky, A., Permat, Y. (in review), Hazard perception differences between young-novice and experienced drivers: A driving simulator study. Submitted to *Hum Factors*.
20. Pollatsek, A., Fisher, D.L., Pradhan, A.K., (2006). Identifying and Remediating Failures of Selective Attention in Younger Drivers. *Curr Dir Psychol Sci*, 15 (5), 255-259.
21. Regan, M.A., Triggs, T.J., Godley, S.T. (2000). Simulator-based evaluation of the DriveSmart novice driver CD-ROM training product. Proc. of the Road Safety Research, Policing and Education Conference
22. Sagberg, F, Bjørnskau, T. (2006). Hazard perception and driving experience among novice drivers. *Accident Anal Prev*, 38(2), 407–414.
23. Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. *Ergonomics*, 50(8), 1235–1249.
24. Wallis, T.S.A., Horswill, M.S. (2007). Using fuzzy signal detection theory to determine why experienced and trained drivers respond faster than novices in a hazard perception test. *Accident Anal Prev*, 39, 1177–1185.
25. Wang, Y.B., Peng, P., Liang, L., Zhang, W., Wu, S. (2007). Road Hazard Reaction Testing Using Driving Simulation: the Novice vs. the Experienced Drivers. Proc. of the 2007 IEEE IEEM.
26. Wang, Y.B., Zhang, W., Salvendy, G. (2010). A Comparative Study of Two Hazard Handling Training Methods for Novice Drivers. *Traffic Inj Prev*, 11(5), 483-491.
27. Yona, O., Borowsky, A., Oron-Gilad, T., Parmet, Y. (2012). The use of a homogeneity measure to identify hazard perception abilities of novices and experienced drivers in a driving simulator. In Proc. of the Human Factors and Ergonomics Society 2012 Annual Meeting 2012. pp. 2266-2270.
28. Zuckerman, M. (1994). *Behavioral expressions and biosocial bases of sensation seeking*. New York: Cambridge Press.