Redevelopment of Victoria's Hazard Perception Test

John Catchpole# and Corinne Leadbeatter~

Senior Research Scientist, ARRB Transport Research Ltd
~ Manager, Driver Assessment Projects, Registration and Licensing Department, VicRoads

THE IMPORTANCE OF HAZARD PERCEPTION SKILLS FOR NOVICE DRIVERS

The over-representation of young and inexperienced drivers in traffic accidents (compared to more experienced drivers) is widely acknowledged as one of the most intractable of safety issues. In highly motorised countries, the rates of accident involvement, injury and death per distance travelled are higher for young novice drivers than for all other segments of the driving population, with the possible exception of the very oldest drivers (say those aged 80 years or more).

More than three decades of research into this issue has revealed that both age-related and experience-related factors contribute to the elevated accident involvement of novice drivers (e.g. Catchpole, Macdonald and Bowland, 1994). Among factors related to lack of driving experience, there is considerable evidence that poor hazard perception skills make an important contribution to novice driver accident involvement. Young, inexperienced drivers are slower than more experienced drivers to detect and respond to hazards in the driving environment (Quimby and Watts, 1981) and less likely to detect child pedestrians and cyclists in the driving environment (Egberink, Lourens and van der Molen, 1986). Slow hazard detection (measured in a driving simulator) is associated with a history of greater self-reported accident involvement (Quimby, Maycock, Carter, Dixon and Wall, 1986). Novice drivers have been found to be particularly over-represented in those types of accidents that are most likely to result from poor hazard perception (Catchpole, Cairney and Macdonald, 1994; Catchpole, 1998).

This paper provides a brief description of the current hazard perception test (HPT) and an overview of the process of developing VicRoads' new updated and expanded HPT.

THE HAZARD PERCEPTION TEST

The HPT was developed between 1989 and 1991 and was implemented throughout Victoria in 1996 as part of the introduction of computerised licence testing. The HPT is currently undertaken at the probationary licence testing stage along with a knowledge test and a practical drive test.

The aim of the HPT is to screen novice drivers on their ability to assess traffic situations and to make safe driving decisions. The existing HPT consists of a computerised touch screen that displays a series of moving traffic scenarios. Test applicants respond to each scene by touching (or abstaining from touching) the screen to indicate when it is safe or necessary to take appropriate action to reduce the risk of a crash in the context of the presented video traffic situation. It attempts to assess the learner's ability to observe the whole road scene; to identify and assess possible hazards in the traffic environment and to make timely responses to avoid or deal with the possible hazard.

One of the difficulties with practical driving tests is the non-standard testing environment. Traffic conditions, the number of incidents and the level of difficulty of incidents vary from applicant to applicant. McKenna and Horswill (1998) described some advantages of hazard perception assessment. These include removing the differences between different examiners; exposure of candidates to varying driving conditions and the ethical problem of exposing drivers to potential accident-producing scenarios. It also assists in encouraging learner drivers to reach a particular standard of hazard perception prior to testing and in setting an agenda for training. Congdon and Cavallo (1999) agree and further state that such a hazard perception test could increase awareness of and importance placed on skills dealing with other road users and the driving environment and encourage novice drivers to gain more supervised driving experience. Deery (1999) points out that the learning of responses to a range of potentially hazardous situations is probably one of the major contributions to driver safety that is acquired through experience.

VicRoads commissioned the Australian Council for Educational Research (ACER) to assess the predictive validity of the current HPT. Congdon (1999) outlined the main findings of the study, which are summarised below.

- The test in its current form has a low reliability, which in part is due to the relatively short length of the test, and partly due to some items eliciting responses that are inconsistent with the responses to other items;
- The HPT measures were evident in identifying fatal and serious injury accident involvement for novice drivers generally; and
- The appearance of the HPT measure as a statistically significant contributor to the prediction of fatal, serious and other injury accidents is a positive outcome for the test. If the hazard perception skills of novice drivers can be measured with greater precision, then the HPT's predictive powers are more likely to increase.

UPDATE AND EXPANSION OF THE HPT

The current HPT was developed at a time when technological constraints limited the test to 12 traffic scenarios each of seven-second duration only. The 12 items seen by each candidate are currently selected from a pool of only 30 hazard perception items. A number of improvements to the HPT are required to make the test easier for licence applicants to comprehend and to maximise the test's precision, reliability and effectiveness. These include:

- a greater range of traffic scenarios to encompass all appropriate high risk situations for novice drivers;
- an increased number of items in the test;
- longer video sequences to better simulate real world driving;
- scripting of scenarios to ensure an even spread of correct response windows across test item types;
- elimination of poor test items;
- review and improvement of test instructions to maximise comprehension for licence applicants;
- review of the scoring methodology (overall test and individual test items); and
- improvement of the video quality and graphic interface design.

Preservation of current HPT features

Research to date suggests a number of features of the current test are critical to its validity and should be preserved during the redevelopment. These include:

- Preserving of the current timed interactive methodology;
- The establishment of "windows" of correct response within the video segment of each item which are of sufficient length and appropriate placement to ensure that the item is not simply assessing motor reaction time; and
- Groups of test items should each have sufficient items to allow random selection of specific items for presentation in a given test. All groups of items should be incorporated in every application of the test using stratified sampling. Analysis of the developed test items needs to examine carefully the contribution of the groups of items and to allow for possible weights for groups as well as for specific items.

THE NEW HAZARD PERCEPTION TEST

Human-Computer Interface

The update of the HPT included the development of a completely new applicant/computer interface utilising a standard mouse to collect candidate responses. In addition, a complete overhaul of the existing touch screen interface was carried out to ensure the greatest possible degree of compatibility between the mouse-response and touch-screen-response interfaces[1]. This involved consideration of design issues, particularly with respect to overall structure, text and language. Further design considerations are consistency of instructions, particularly for navigation and item responses. Navigation for the package is generally limited to a single instruction to reduce confusion or distraction. The aim of the design elements is to keep the user's focus as much as possible on the information that they are seeking by minimising the difficulty of navigating through the package.

[1] Interface design work, including a review of usability problems in the current interface, was carried out for the project team by Prof. Gitte Lindgaard, then of Swinburne University of Technology and presently Chair, User-Centred Design, Dept. of Psychology, Carleton University, Ottawa, Canada.

Potential useability problems were identified by means of an heuristic evaluation of the current HPT. The purpose was to identify the merits and shortcomings of the user interface. Usability problems identified in the existing HPT mainly concerned the instructions, inconsistent use of colour, and screen design issues. These were confirmed by usage observations, a usability evaluation and interviews with applicants and stakeholders.

The updated HPT takes candidates through a session comprising three parts:

1. Instructions;
2. Practice items; and
3. Test items

Progress through the HPT is controlled by the candidate, who can choose to repeat the instruction section as many times as desired before proceeding to the practice items, and can repeat the practice items before proceeding to the test items.

Instructions

The update of the test also included a review of the instructions and feedback messages read and heard by test candidates[2]. The level of language used throughout the HPT package has been kept as simple as possible to cater for the widest possible range of users. As a general guide, the complexity of the text was aimed at a comfortable level for an average 12-year-old.

In each item of the updated HPT, a traffic scenario is presented to test candidates in a short moving video clip on the computer screen. Each test item asks the candidate to indicate when it would be appropriate to perform one of four driving actions:

1. slow down,
2. overtake,
3. make your turn, or
4. move off.

Feedback messages

After each test item, the candidate receives an immediate feedback message telling him or her whether or not a response was registered. In the mouse-response version of the interface, the feedback message states either "You have clicked a mouse button" or "The video has finished". The latter message tells candidates that no mouse response was registered. For practice items but not test items, the candidate receives a second feedback message concerning the correctness of their response. This message can take a variety of forms, including "You clicked, but too early", "You clicked, but too late", "You should not have clicked for this item" and several others. This type of feedback helps candidates to decide whether they have correctly understood the task and whether they need to repeat the practice items before proceeding to the test items.

TRIAL OF THE USER-FRIENDLINESS OF THE INTERFACE AND INSTRUCTIONS

A trial of the new interface and instruction set was conducted to allow any usability problems that may exist to be identified and eliminated prior to the trialling and psychometric assessment of new test items. Thus, any problems that may be identified during later trialling of new items will most likely relate to the items rather than being the result of usability problems with the new interface.

Audio recordings were made of all instruction and feedback messages. New software was developed to deliver the new interface, instructions and feedback. These new materials were combined with software modules and item video clips from the current HPT to produce the high-fidelity prototype used in the trial.

Ten people participated in the usability trial in individual video taped sessions. Each participant was interviewed after the test session to identify any problems they might have experienced during the trial.

The post-test interviews revealed that participants were, by and large, quite enthusiastic about the prototype hazard perception test. In particular, participants commented favourably on:

---

[2] The review and update of the instruction set and feedback messages was carried out by Peter Congdon and Brian Doig (Australian Council for Educational Research) with the assistance of other members of the project team.

- the usefulness of practice items for understanding what to do in the test;
- the ability to repeat the practice items;
- the still preview screen that preceded each moving video;
- the clear instructions and pictures; and
- the speedometer indicating the vehicle's current speed.

While no serious usability problems were identified, a number of minor problems emerged. The most important of these was that many participants clicked the mouse button when not required to do so, sometimes in an attempt to control the pace of presentation at times when the pace was under automated control. However, these additional clicks apparently did not interfere with participant performance on the HPT. Unnecessary mouse clicks will be monitored during the item trial to confirm this conclusion.

Following the usability trial, a number of minor changes were made to the interface and to the wording of some instructions. At several key points in the session, timing was changed from automated to user-controlled.

ANALYSIS OF NOVICE DRIVER ACCIDENTS

It is important that the traffic scenarios developed for the test are based on the types of situations leading to accidents in which young and inexperienced drivers are most likely to be involved. An analysis of novice driver accident data was therefore conducted prior to development of the test items to determine the types of hazardous situations that should be depicted in the traffic scenarios.

Drivers aged 18 to 59 years involved in casualty accidents reported to police in Victoria during the years 1986 to 1997 were included in the analysis. Drivers of trucks, buses and motorcycles and drivers with a blood alcohol concentration in excess of 0.020 g/100 mL were excluded. After all exclusions, over 247,000 accident-involved drivers remained in the analysis; in many cases, this included more than one driver from the same accident.

Each accident-involved driver was classified according to the VicRoads Definitions for Classifying Accidents (DCA codes) and the role of the driver in the accident. For example, drivers involved in rear end accidents (DCA 130) were given the involvement code 130-1 (driver of the rear vehicle in a rear end accident) or 130-2 (driver of the front vehicle); separate codes were also allocated to the small minority of drivers not involved in the initial collision or whose role in the collision was not known. This classification system yielded a total of 306 different types of accident involvement, of which 89 involvement types accounted for 96 per cent of all accident involvements. Using this classification system, the most frequently occurring types of accident involvement for drivers with less than one year of driving experience since acquiring a probationary licence were found to be as shown in Figure 1. The 17 highest ranked involvement types shown in Figure 1 accounted for 70 per cent of all accident involvements of drivers in their first year since acquiring a solo licence.

The highest frequency accident involvement types, as shown in Figure 1, reflect not only the skill deficits of the drivers involved but also the most commonly occurring conflict situations in normal driving. For this reason, there is a degree of commonality between the accident involvement type rankings for drivers of different ages. For example, the seven most frequent involvement types for first year drivers also fill six of the top seven places for all experience levels combined and five of the top seven places for drivers with 15 or more years of experience. To focus specifically on the skill deficits of inexperienced drivers, further analysis was conducted to identify those involvement types in which first year drivers are most over-represented by comparison with more experienced drivers – that is, the involvement types that form a much greater proportion of all accident involvements for first year drivers than they do for experienced drivers. The most over-represented accident involvement types for first year drivers are shown in Figure 2.

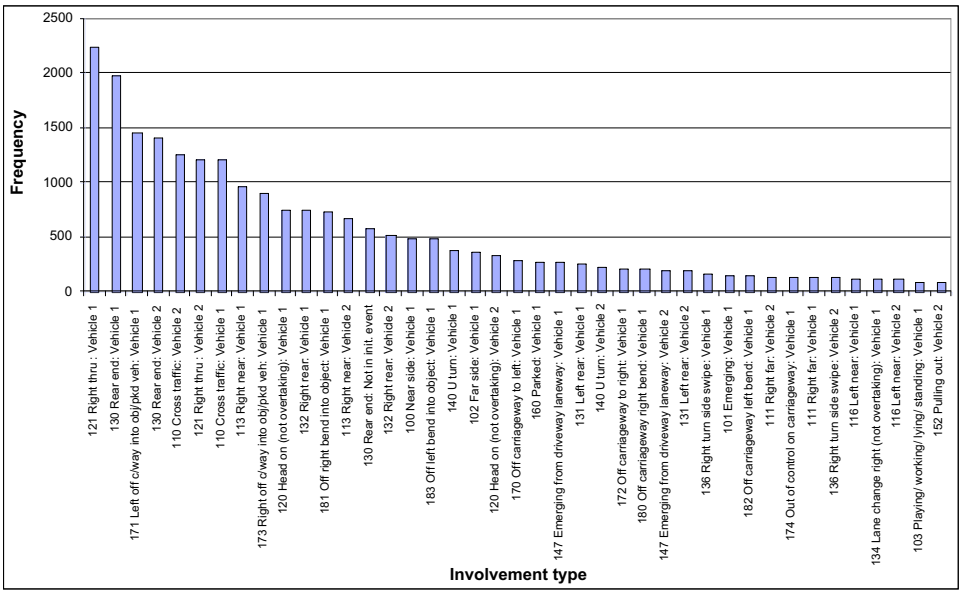

Figure 1. Highest frequency accident involvement types for first year drivers, Victoria, 1986-97.

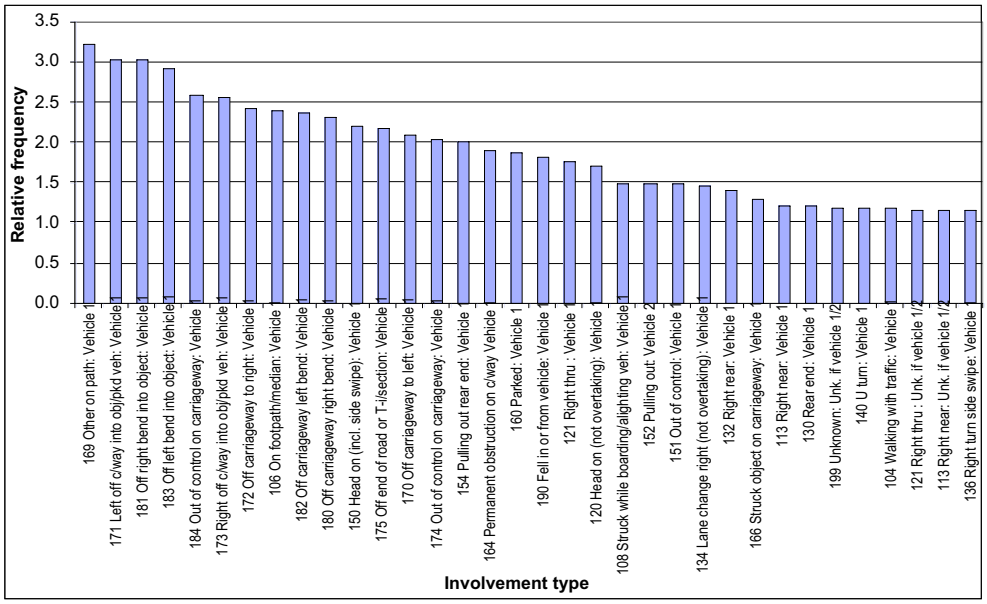

Figure 2. Most over-represented accident involvement types for first year drivers, Victoria, 1986-97.

DEVELOPMENT OF TRAFFIC SCENARIOS

In each test item, a traffic scenario is presented to test candidates in a short video clip on the computer screen. Each video clip is scripted in great detail to ensure that the critical stimuli will be visible to the test candidate at the appropriate time. Preparation of the video scripts for the test items was undertaken in four stages.

In the first stage, the accident involvement types to be depicted in the traffic scenarios were selected. A starting point for this selection was the list of accident involvement types previously identified as being the most frequently occurring or the most over-represented for first year drivers. The selection process also took into account novice driver skill deficits identified by previous research; the involvement types represented by those items from the existing HPT with the highest predictive validity; the degree of likelihood that hazard perception skill problems contribute to each accident involvement type; and the technical limitations of the hardware platform used to present test items to candidates. A total of 31 accident involvement types, including pedestrian, single vehicle and multi-vehicle types, were selected for representation in test items.

The second stage comprised the development of 49 traffic scenario outlines that included or culminated in the appearance of the hazard or conflict appropriate to each of the selected accident involvement types. Some involvement types were addressed by more than one scenario outline. Each scenario outline was essentially a specification of the movements of the various road users involved in the lead up to the conflict or hazard. The sequence of road user movements described in the scenario outlines was based mainly on experience gained in previous research into the road user behaviours preceding accidents of various types (Catchpole, Cairney and Macdonald, 1994; Cairney and Catchpole, 1991). Additional information was obtained from police accident report form diagrams and narratives as required.

In the third stage, several variants of each scenario outline were constructed by varying the number of road users involved; the precise timing of the movements of the road users; the type of vehicles involved; the age and sex of pedestrians and cyclists; the type of location; and light and weather conditions. The specification of environmental conditions for each variant was guided in part by the typical environments for each accident involvement type, as identified in the earlier analysis of the accident involvements of first year drivers. A total of 150 scenario variants was constructed.

Finally, a detailed item script was developed from each of the 150 scenario outlines, specifying the precise speed and timing of all road user movements and all details required by the filming contractor to produce the video clips for the 150 test items.

All test items can be classified as either *response* items or *no response* items. *Response* items have a time period (called the 'correct response window') when the candidate should make a response; that is, a time when it is necessary or safe to perform the specified driving action. *No response* items do not have a correct response window; there is no time when it is necessary or safe to perform the specified driving action.

Two critical considerations during the generation of the scenario variants and the final scripts was the placement within the item of the 'correct response window' and the balance of *response* versus *no response* items. Experience with the current HPT has shown that some candidates have been advised to respond according to strategies that have no relation to the road user movements depicted in the video clip. The longer video clips to be used for the new HPT items (15 seconds in the new HPT, compared to 7 seconds for most of the items of the existing test), the much larger item pool (expected to be in the vicinity of 100 items) and the greater length of the updated test (increasing from 12 to approximately 30 items for each candidate) provide much more opportunity to thwart any attempt to pass the test by using this type of strategy. For items of each different type (based on the instructions to the candidate), there is a broad distribution of correct response window locations and a proportion of *no response* items, so that no simple strategy will yield more than a chance proportion of correct responses.

THE NEXT STEPS

Filming

As the scripts and other specifications for filming are based on rigorous analysis of traffic hazards and psychometric testing considerations, no creative interpretation will be permitted in the production of video material. There are three main stages involved with the filming of the scenarios:

Stage 1 - Filming or video recording of 150 traffic scenarios according to scripts and other defined specifications;

Stage 2 - Editing of raw material as directed by VicRoads to combine driver perspectives to give the impression of a single continuous sequence of driver activity; and

Stage 3 - Conversion of raw material to MPEG2 format, including compression for optimised delivery in a defined PC environment.

All 150 traffic scenarios require the recording of a driver's-eye perspective as may be expected when looking ahead through a vehicle windscreen. Some scenarios also require driver's-eye perspectives as may be expected when looking to the right or left. It is expected that the duration of material to be recorded for each traffic scenario will be approximately 30 to 40 seconds, of which about 15 seconds will be centred on the specific scenario to be used in construction of eventual test items.

Filming and review of test sequences will permit fine-tuning of technical solutions and will evaluate the suitability of specific filming requirements. These include camera locations; camera field-of-view and line of

sight for various driving situations; suitability of camera and/or recorded image stabilisation techniques; depiction of scripted speeds when viewed on the delivery platform; distances at which indicator and brake lights are visible in day-light conditions; and minimal clarity required for cues to be reliably detected by HPT candidates.

It is important for HPT items to depict a broad range of environmental conditions so that learner drivers see what they learn during their preparation for the HPT as relevant to real life driving. The scripts specify the events and objects that are critical either to the interactions of the scripted road users or to the plausibility of the traffic scene, but still allow some flexibility to the filming contractor in relation to non-critical events. For example, for many of the scripts the precise characteristics of the person(s) involved are not critical. However, over the 150 scripts, it is essential that a broad range of characteristics be depicted which are representative of the Victorian community. This includes non-stereotypical gender and cultural representation and inclusion of a range of ages, mobility and vehicles. It is also a requirement that scripts filmed in wet conditions not exceed 20 per cent of all scripts filmed.

Some of the test items will require the candidate to make a judgement about speed and gap selection. To enable such judgements to be made, the field of view must be such that it results in a realistic impression of speed – that is that the filmed speed gives the impression that the car is moving at the speed shown on the simulated speedometer.

The safe and appropriate management of traffic during filming is a critical element of this project. Safety and traffic management plans will be prepared for all locations where filming is to take place to ensure that actors, film crew and members of the travelling public are not exposed to any inappropriate risks during filming. Traffic management requirements will vary between locations and times of day, with warning signs, lane closures, temporary diversions and other measures being used as appropriate.

Item trialling and test construction

The main purpose of the item trial will be to measure the difficulty and reliability of the newly developed test items by obtaining a substantial body of item response data from a representative sample of the target population. A group of "expert" drivers (highly experienced and with excellent safety records) will also be included in the trial to validate the correct response windows locations nominated for each item by the development team. Following the trial, it is expected that the poorest performing items (those with very low reliability or excessively high or low difficulty) will be dropped from the item pool.

Both mouse-response and touch-screen-response interfaces will be used during the trial. Responses collected using the two interfaces will be compared to determine whether any adjustment of correct response window locations is needed to compensate for differences between the two response mechanisms. The overall performance of the items under the two interfaces will also be examined to determine whether hazard perception testing can be moved to a standard hardware platform (i.e response via mouse) or whether it is necessary to continue with more expensive touch screen hardware similar to that used in the current HPT.

Final test construction will be based on psychometric assessment of the items derived from response data collected during the trial. Test forms will be of equivalent difficulty and will be constructed to include the full range of item tasks (click when you would slow down/overtake/etc.), correct response window locations and accident involvement types represented in the item pool. Test construction will include calculation of raw scores corresponding to various nominal pass rates for the overall test.

Implementation in Victoria's Computerised Licence Testing system

A complete redevelopment of Victoria's Computerised Licence Testing (CLT) system has recently commenced. As part of this redevelopment, the new HPT will be implemented using the latest and most appropriate software technologies to run on high quality, up-to-date hardware platforms. Following full testing, the new CLT and HPT are expected to come into use for licence testing in VicRoads offices in 2001.

CONCLUSION

The update and expansion of Victoria's Hazard Perception Test will lead to a new test that is expected to measure the hazard perception skills of novice drivers with greater precision, thereby improving the test's

predictive powers. The new HPT is expected to be easier for licence applicants to comprehend and to have improved reliability and effectiveness. Overall, the updated test is expected to provide:

- improved predictive validity;
- a greater range and number of traffic scenarios;
- longer video sequences to better simulate real world driving;
- an even spread of correct response windows and window placement across test item types;
- no scope for successful response strategies based on information other than the traffic scenario depicted;
- improved video quality and graphic interface design; and
- greater public acceptance.

REFERENCES

Cairney P and Catchpole J (1991). *Road user behaviours which contribute to accidents at urban arterial/local intersections.* Research Report ARR 197. (Australian Road Research Board: Vermont South, Victoria)

Catchpole J (1998). Age, driving experience and type of accident involvement. In: *Road Safety Conference 1998: Proceedings*, volume 1, pp 196-201. (Land Transport Safety Authority: Wellington, NZ)

Catchpole J, Cairney P and Macdonald W (1994). *Why Are Young Drivers Over-Represented in Traffic Accidents?* Special Report No. SR 50. (Australian Road Research Board: Vermont South, Vic).

Catchpole J, Macdonald W and Bowland l. (1994). Young driver research program: The influence of age-related and experience-related factors on reported driving behaviour and crashes. Report No. CR 143. (Federal Office of Road Safety: Canberra, ACT)

Congdon, P (1999). *VicRoads Hazard Perception Test, Can it Predict Accidents?* Australian Council for Educational Research, Victoria, Australia.

Congdon, P & Cavallo, A (1999). Validation of the Victorian Hazard Perception Test. In: *Research Policing Education Road Safety Conference.* (ATSB: Canberra, ACT)

Deery H (1999). Hazard and risk perception among young novice drivers. *Journal of Safety Research, 30(4)*, 225-236.

Egberink HO, Lourens PF and van der Molen HH (1986). Driving strategies among younger and older drivers when encountering children. *Accident Analysis and Prevention, 18 (4)*, 315-324.

McKenna, F, and Horswill, M (1998). Hazard Perception and its relevance for driver licensing. Issues on Driver Training and Licensing. *IATSS Research, 23* (1) 36-41.

Quimby AR, Maycock G, Carter ID, Dixon R and Wall JG (1986). *Perceptual Abilities of Accident Involved Drivers.* Research Report RR 27. (Transport and Road Research Laboratory: Crowthorne, Berkshire, UK).

Quimby AR and Watts GR (1981). *Human Factors and Driving Performance.* Laboratory Report LR 1004. (Transport and Road Research Laboratory: Crowthorne, Berkshire, UK).