Title: Development and validation of the Spanish Hazard Perception Test.

Authors: Castro\textsuperscript{a}, C., Padilla\textsuperscript{a}, J. L., Roca\textsuperscript{b}, J., Benítez\textsuperscript{c}, I, García-Fernández\textsuperscript{d}, P., Estévez\textsuperscript{e}, B., López-Ramón\textsuperscript{c}, M.F., and Crundall\textsuperscript{d}, D.

Affiliations:

\textsuperscript{a}Cimcyc. Mind, Brain and Behavior Research Center. Faculty of Psychology. University of Granada. SPAIN

\textsuperscript{b}Faculty of Psychology. University of Valencia. SPAIN

\textsuperscript{c}Faculty of Education. Teruel. University of Zaragoza. SPAIN

\textsuperscript{d}Nottingham Trent University. UK

Keywords: Hazard Perception, novice drivers, experienced drivers, driving offenders, non-offenders, driving safely.

Corresponding author: Candida Castro Email address: candida@ugr.es

Faculty of Psychology. Cimcyc – Mind, Brain and Behaviour Research Centre


Phone: + 34 958 240663 +34 655 99 40 59 Fax: +34 958 246239

DOI: 10.1080/15389588.2013.879125
Driving Hazard Perception Test in a Spanish Setting

ABSTRACT

Objective: The aim of the current study is to develop and obtain validity evidence for a Hazard Perception test suitable for the Spanish driving population. To obtain validity evidence to support the use of the test, the effect of hazardous and quasi-hazardous situations on the participants’ Hazard Prediction is analysed and the pattern of results of drivers of different driving experience: learner, novice and expert drivers and re-offender vs. non-offender drivers, is compared. Potentially hazardous situations are those that develop without involving any real hazard (i.e., the driver didn’t actually have to decelerate or make any evasive manoeuvre to avoid a potential collision). The current study analysed multiple offender drivers attending compulsory re-education programmes as a result of reaching the maximum number of penalty points on their driving licence, due to repeated violations of traffic laws.

Method: A new video-based hazard perception test was developed, using a total of 20 hazardous situation videos plus 8 quasi-hazardous situation videos. They were selected from 167 recordings of natural hazards in real Spanish driving settings.

Results: The test showed adequate psychometric properties and evidence of validity, distinguishing between different types of drivers. Psychometric results confirm a final version of the hazard perception test composed of 11 video clips of hazards and 6 video clips of quasi-hazards, for which an overall Cronbach’s alpha coefficient of .77 was obtained. A lack of ability to detect quasi-hazards and distinguish them from hazardous situations was also found for learner, novice and re-offender drivers. Learner drivers obtained lower average scores than novice and experienced drivers with the hazardous situation videos; and learner drivers obtained lower average scores than experienced drivers with the quasi-hazardous situation videos, suggesting that the ability to correctly identify hazardous traffic situations may develop early by accumulating initial driving experience. However, the ability to correctly identify quasi-hazardous situations may develop later with the accumulation of further driving experience. Developing this ability is also difficult for re-offender drivers.
Conclusion: The test has adequate psychometric properties and is useful in distinguishing between learner, novice and expert drivers. In addition, it is useful in that it analyses the performance of both safe and unsafe drivers (re-offenders who have already lost their driving licence).
INTRODUCTION

Hazard perception is the skill of detecting, evaluating and reacting to events on the road that have a high probability of producing a collision (Crundall et al. 2012). The skill of hazard perception has been reported to be a specific driving ability that correlates with the risk of having an accident (Darby, Murray and Raeside, 2009). Hazard detection tests measure the ability to identify hazardous situations during the activity of driving. These tests can be used as a diagnostic tool and a training tool to improve the skills of those drivers who lack the ability to detect hazards.

According to McKenna and Crick (1991) and Mills et al. (1998), hazard perception can be considered as the ability to “read” the road. That is, the ability to interpret the key information from the driving environment that allows the driver to anticipate what is going to happen. McKenna and Crick (1991) emphasised that a hazardous driving situation is one where it is necessary to perform an evasive action, for instance, braking or swerving abruptly.

Horswill and McKenna (2004) considered hazard perception as the “situation awareness” (Endsley, 1995) of potentially hazardous incidents in driving. Consequently, hazard perception would involve not only perceiving the hazard but also understanding the situation and anticipating what will happen in that situation. That is, being able to predict and anticipate future traffic situations in order to make the right decisions and perform the correct manoeuvres, using previous knowledge and the clues provided by the situation.

The hazard detection task

Interest in developing hazard perception tests is growing in numerous countries, for example in the U.S. (Chen, Mourant and Nie, 2011), and in Israel (Borowsky et al. 2010, 2013, Meir et al. 2014). In a traditional test of Hazard Perception (e.g., Wetton, Hill and Horswill, 2011), the driver is asked to perform a computer task in which videos of driving scenes filmed from the perspective of the driver of a moving vehicle are shown. The participants are told they must detect certain hazardous events (for instance, a vehicle emerging from one side at a crossroads) by pressing a key as soon as the danger is noticed. McKenna and Crick (1991) found that expert drivers detected the danger before
Driving Hazard Perception Test in a Spanish Setting

novice drivers. However, other authors (Sagberg and Bjørnskau, 2006) using the standard methodology of Hazard Perception did not find significant differences between novice and expert drivers. If expert drivers use predictive cues to allow them to anticipate the hazard in advance, it is possible that these cues were not present in the test clips used by Sagberg and Bjørnskau (2006). If the content of these videos did not contain predictive elements then the videos may have been evoking simple reactions to abrupt events rather than encouraging participants to perform a proactive search that could prime the hazard prediction.

Young people have faster reaction times that could mask the effect of experience. If the task requires pressing a button in response to the presence of a stimulus (an unpredictable hazard) that appears suddenly, it is possible that young people show faster reflexes that could mask the advantage of experience shown by expert drivers (Jackson et al. 2009).

The hazard prediction task: What Happens Next?

Contrary to the classic Hazard Perception Test, the Hazard Prediction Task, What Happens Next? (WHN) not only explores the perception of danger, but also the driver’s understanding of the traffic scene. From these (perception and comprehension), it is possible to plan and predict what will happen next. A correct prediction of the danger is the determining factor for making appropriate decisions about the manoeuvres to be performed. These three stages: perception, comprehension and projection from Endsley’s (1995) Situation Awareness model are essential for hazard prediction (Horswill and McKenna, 2004; Jackson, et al. 2009).

Situation Awareness (SA) is a mental representation that involves the comprehension of objects, events, people, their interactions, environmental conditions and any other type of factor relating to a specific situation that could affect the development of human tasks, complex or dynamic. When we have “situation awareness”, we know what is happening and we can plan what must be done. In addition, it can be defined as that which is necessary in order not to be surprised. With regard to hazard perception, Level 3 SA equates to the ability to answer the question “WHN?” by detailing the most relevant event that is about to occur. For instance, one might predict that a pedestrian seen exiting a house and disappearing behind a parked vehicle may emerge from behind
the vehicle and step into the road. This presumably shows a good level of situational awareness that has initially perceived the pedestrian, comprehended the possible hazards that the pedestrian could pose, and then predicted the most likely combined trajectories of both the driver and pedestrian to realise that a hazard might appear. A failure at any of these levels of SA could result in a failure to avoid the hazard.

According to Jackson et al. (2009) and Crundall et al. (2012), there are some conditions in the videos that make them more or less easily distinguished by different groups of drivers, novices vs. experts, or unsafe vs. safe drivers. Different manipulations of the videos were carried out, in which researchers included at least some videos with an environmental factor associated with previous knowledge that might facilitate hazard detection. For instance, these authors used Environmental Prediction (EP) hazards, situations where the precursor of the hazard and the hazard were not the same, the precursor being considered a clue from the environment that could help to predict the hazard. EP hazards tend to be occluded hazards, where a pedestrian is hidden by a parked vehicle, a broken-down vehicle is hidden by a blind bend or an emerging vehicle is masked by another blocking vehicle. In these cases the driver must consider whether the parked vehicle, blind bend, or non-hazardous moving vehicle is likely to be hiding a hazard. Thus, they are predicting the possibility of a hazard on the basis of the statistical likelihood that an aspect of the environment may hide a hazard.

Behavioural Prediction (BP) hazards do contain a precursor, but the link between the two is more direct. If a pedestrian is visibly walking on the pavement and then turns to step into the road, then they are both precursor and hazard. They are the precursor while walking along the pavement, perhaps looking over their shoulder, but they become the hazard when they step into the road. Thus, one uses the behaviour of the pedestrian (walking along, looking over their shoulder, etc.) to predict whether they will become a hazard (hence BP).

**Repeat offenders**

Did multiple offenders obtain different hazard perception scores from normal/safe drivers? (see the classic study by Pelz and Krupat, 1974). The implication is that good drivers are more likely to avoid accidents than are drivers with multiple driving violations. And therefore, accident
Driving Hazard Perception Test in a Spanish Setting

prevention or recidivism prevention courses for repeat traffic offences could be worth (McKnight and Tippetts, 1996). According to Simon and Corbett (1996), the results of drivers’ accident history are positively related to offending, for example, an accident history of speeding and driving under the influence of alcohol. Road safety researchers have examined the relationship between sensation-seeking and risky behaviour (e.g. driving while impaired, speeding, following too closely) as well as its consequences (e.g. collisions, violations) (see Jonah’s, 1997 review).

Taylor, Baruya and Kennedy (2002) stated that accident frequency for all categories of accident increased rapidly with mean speed – the total injury accident frequency increased with speed to the power of approximately 2.5, thus indicating that a 10% increase in mean speed results in a 26% increase in the frequency of all injury accidents. In addition, Laapoti et al. (2001) reported that the number of accidents and offences was highest among young males, whose accidents took place later at night than female or older drivers’ accidents.

Lapham et al. (2006) stated that repeat offenders are more likely than DUI (first time driving under the influence) to be involved in fatal motor vehicle crashes, hit and run collisions with pedestrian fatalities and to have a high blood alcohol concentration (BAC) when driving (0.15 mg/dl and above) (Beirness et al. 1991; Fell, 1993, 1995; Solnick and Henneman, 1994).

As an European example, Yahya and Hammarstroöm’s (2011) recent report describes the results of a study in Sweden, a total survey, conducted with the aim of determining to what extent differences exist between drunk and sober drivers as regards the penalty for traffic offences that result in road accidents where people were killed or injured. There were a total of 1,995 records in the register for the specified group, the main group, 88% of whom were men and 12% women. The proportion of drunk drivers was 25% for men and 20% for women. The group of drivers below the age of 35 was over-represented in the register in relation to the general population. In the main group, 31% had committed an offence during the previous 10 years. Driving offences had been committed previously by 9%. More than three previous offences had been committed by 14%. The proportion of women with previous legal proceedings was smaller when compared to the main group. One in five of those convicted in a district court for “causing another person’s death” and “causing bodily
Driving Hazard Perception Test in a Spanish Setting

harm or illness” had been drunk. Of those convicted for “creating danger to another person”, two thirds had been drunk.

Research Aims

The objective of the current work was twofold: (1) develop a hazard perception test suitable for the Spanish driving population. And (2) to obtain validity evidence to support the use of the test with Spanish drivers.

To achieve the first aim, a new video-based hazard perception test was developed based on the design used by Jackson et al. (2009), which requires participants to predict what will happen next after the screen turns blank immediately prior to a hazardous event. Drivers are asked three questions after each video clip of a hazard: What the hazard is?, Where the hazard is?, and WHEN?, making it possible to analyse the detection of the hazard as well as the driver’s situation awareness.

To obtain validity evidence to support the use of the test, the differences in hazard perception scores between groups of drivers varying in experience (learner, novice and experienced drivers) and in driving recidivism (multiple offenders and non-recidivist drivers) were checked. According to previous research (Crundall et al. 2010, 2012), learner drivers would be expected to obtain the lowest hazard perception scores as compared to novice or more experienced drivers. Novice drivers have restricted scan patterns (e.g. Crundall and Underwood, 1998), they have longer fixation times on road-relevant objects (Chapman and Underwood, 1998), reduced peripheral attention, (Crundall et al. 2002), a poorer lane maintenance (Land and Horwood, 1995), and are slower to respond to hazards (McKenna and Crick, 1991).

Secondly, the pattern of results of re-offender vs. non-offender drivers was also explored. The current study analysed multiple offender drivers attending compulsory re-education programmes as a result of reaching the maximum number of penalty points on their driving licence, due to repeated violations of traffic laws. Despite the usefulness of studying the performance of unsafe drivers in order to ascertain their driving behaviour and help improve it, such a sample has not been analysed in recent published works and scarcely at all in the past. Further analysis is needed in the future.
Finally, to provide a complementary strand of validity evidence the participants’ performance in response to the hazard perception video clips (which, by definition, always ended prior to an actual hazardous driving situation), and to the quasi-hazardous video clips, in which the potentially hazardous situation developed without involving any real hazard was also analysed.

METHODS
Participants
A total of 73 drivers took part in the current study. Most of them (47) were males. Three groups were considered (see also Table 1): (a) 14 learner drivers (18-36 years, mean age = 21.7, St. Dev. = 5.8), who were taking lessons to obtain a driving licence for the first time and were recruited from either a collaborating driving school in Granada (Spain) or the School of Psychology of the University of Granada; (b) 16 novice drivers (18-29 years, mean age = 20.6, St. Dev. = 3.5), who were in possession of a driving licence and had less than two years’ driving experience, recruited from the School of Psychology of the University of Granada (Spain); and (c) 14 experienced drivers (28-60 years, mean age = 37.5, St. Dev. = 9.0), who possessed a driving licence and had more than eight years’ driving experience, and were recruited from the driving population in Granada (Spain). No driver in any of these three age groups had previously lost all the points on their driving licence as a consequence of committing repeat driving offences, and thus they were considered non-recidivist drivers.

In addition, a group of 29 multiple driving offenders was recruited from a collaborating driving school in Granada (Spain) while they were attending a driving re-education and recidivism prevention course (i.e., the course known as “Recuperación Total”, which is compulsory is Spain for drivers who have been banned from driving after losing all the available points on their driving licence). Since we planned to compare this group of multiple offenders to the previously described non-recidivist drivers (excluding learners, who are not actual drivers), the multiple offender group
Driving Hazard Perception Test in a Spanish Setting

was composed of 15 novice drivers (19-28 years, mean age = 23.2, St. Dev. = 2.7) and 14 experienced drivers (28-53 years, mean age = 38.0, St. Dev. = 7.7) (see also Table 2).

The experimenters in the current study adhered to the ethical principles set out in the World Medical Association Declaration of Helsinki for research involving human participants.

Development of the Spanish Hazard Perception test

A video-based hazard perception test was developed to be applied to the Spanish driving population. The initial version of the test included 20 hazard perception situations and 8 quasi-hazard perception situations that were recorded from a position approximate to the driver’s point of view. The clips showed real traffic situations in which a potentially hazardous event developed when the driver approached. All hazard perception clips were obtained during genuine and unaltered driving sessions and the actual driver had to react before the hazard to avoid a collision (for example, by slowing down or by making an evasive manoeuvre). The video clips lasted between 7 and 52 seconds and stopped just before the actual hazard unfolded. When the video clips stopped, a blank screen was shown to avoid the participants seeing any more of the traffic scene. Then, after each video clip, the participants had to answer three questions: (a) What is the hazard? (b) Where does the hazard appear? and (c) What is going to happen in the traffic scene after the video stops? (i.e., WHN?). The first and the third questions required a short written answer and the participants were given a brief space in the questionnaire dossier to respond. The second question required participants to choose one of three alternatives: “in the left part of the scene”, “in the right part of the scene” or “in the middle of the scene”. Each question allowed the participant to get 2 points for a correct response, 1 point for a partially correct response and 0 points for an incorrect response.

In addition to the clips used in the hazard perception test, the participants were also shown 8 clips of quasi-hazardous situations. These video clips were similar to those used in the hazard perception test, with the difference that while the driving session was being recorded, the potentially
hazardous situation developed without involving any final hazard (i.e., the driver didn’t actually have
to decelerate or make any evasive manoeuvre to avoid a potential collision). Clips of hazardous and
quasi-hazardous situations were presented in a randomly pre-determined order within the same block,
and the participants were not informed of the existence of different kinds of clip. All the clips stopped
before the potentially hazardous situation unfolded and were followed by a blank final scene.
Participants were then required to answer the same three questions (although it should be noted that a
correct response for quasi-hazardous situations involved answering by way of response to the third
question that the situation would not develop into an actual hazard).

Each question (i.e., What the hazard is?, Where it is?, and WHN?) allowed the participant to
get 2 points for a correct response, 1 point for a partially correct response and 0 points for an incorrect
response. Therefore, it was possible to obtain a maximum score of 6 points per video clip. Answers
to the test questions could be scored as wrong, partially correct or correct. For instance when
correcting the question ‘What is the hazard?’ participants obtained 2 points if they gave an exact
description of the hazard (e.g. red car in the left lane), 1 point if they gave a partially correct answer
(e.g. a car on the left, but without giving any details of its characteristics or location or if there was
more than one car) and 0 points for an incorrect answer. To correct the ‘WHN?’ question, 2 points
were given if an exact description of what would happen was provided: “The car will give way
because it cannot change lane and allow the red car to merge smoothly with the traffic”, 1 point if the
description was not complete, such as “The car will give way” and 0 points if the answer was
incorrect, i.e. “the car will keep going at the same speed or in the same trajectory, or the car will speed
up”.

For recording the videos, the protocol developed by the University of Nottingham, UK was
used in order to control for the bias involved in recording (see Wetton et al. 2011). All the videos
were preserved in their original versions and were not not retouched. There were no accidents during
the recordings. Ethical principles laid down in the declaration of Helsinki for research involving
human participants were followed in the current study.
To obtain the set of hazardous and quasi-hazardous situation video clips used in the current study, we followed a similar procedure to the one previously used by the University of Nottingham (Jackson et al. 2009) and considered the general principles proposed by Wetton et al. (2011). First, a pool of 167 video clips was recorded using a CANON HD LEGRIA HFR16 digital camera mounted on the vehicle while it was driven across the province of Granada (Andalucía, Spain) between September and October 2012 by two expert drivers who are part of the research team. Twenty-eight hazard perception clips and 8 quasi-hazardous situation clips were then extracted from the video pool. The selection of clips was made in accordance with the recommendations of Wetton et al. (2011). Three external experts (two police officers and a professional driver) were consulted to assess the hazardousness of the different traffic situations. Additionally, the official accident statistics for Spain (Directorate General for Traffic, 2012) were considered in order to select video clips in approximate proportions corresponding to the different types of road (urban, highway or motorway), different potential hazard sources (car, motorbike, pedestrian, truck or bus) and different potential hazard locations (left, right or centre). The videos included different road types, 11% motorway outside the city (A-44 Granada-Jaén and A-92 Málaga-Guadix, highways) and 89% of urban roads (the neighbourhoods of Sacromonte, Albaycín, Almanjayar, Plaza de Toros and Zaidín, which are typical of Andalusia). Hazard situations comprised cars 50%, pedestrians 25%, motorcycles 7%, trucks 11% and buses 7%, and appeared out of side streets, at junctions etc. Furthermore, different types of hazardous traffic situation were included according to the taxonomy proposed by Crundall et al. (2010, 2012). Specifically, in half of the clips there was a noticeable precursor in the traffic environment that might allow the driver to anticipate the hazard source (Environmental Prediction situations, EP), whereas in the other half of the clips there was no precursor and the behaviour of the hazard source turned into a hazard abruptly (Behavioural Prediction situations, BP).

Next, raw video data of selected clips was processed using specialized software (i.e., VLC Media Player, WinX HD Video Converter Deluxe and Adobe Premier Pro CS5). Video processing was aimed only at stopping each clip before the potential hazard unfolded, inserting the blank final
scene in each clip and generating a single video sequence with all the clips to be shown to the participants during data collection sessions.

Finally, the participant dossier also included a socio-demographic questionnaire to gather information such as age, gender, level of education, years driving regularly, years since obtaining different types of driving licence, driving frequency, kilometres driven during the last 12 months, involvement in different kind of traffic incidents during the last 12 months (i.e., accidents with material damage, accidents with victims, “quasi-accidents”, incidents reported to an insurance company and driving fines received) and the number of times they had lost their driving licence, among others (see Tables 1 and 2).

Procedure

Participants completed the test in group sessions. They were recruited from either the School of Psychology of the University of Granada or different collaborating driving schools in Granada (Spain): Autoescuela La Victoria and Autoescuela Luna. First, the participants filled in the socio-demographic questionnaire individually. Then they were presented with the video clips in groups while seated at a distance of between 3 and 5 metres from a projection screen. The clips were displayed on the screen at a 1920 x 1080 resolution using a video projector connected to a standard computer. After each clip, participants were given enough time to answer the corresponding questions. A sequence of all hazardous and quasi-hazardous situation clips was generated in a pseudo-random order and this sequence was counterbalanced per group session.

Data analyses

Once data collection was completed, the participants’ answers were assessed by a single evaluator. To confirm the consistency of the scoring, 15% of the questionnaires, randomly selected from the whole sample, were independently assessed by a second evaluator and the degree of agreement between the two evaluators was assessed by means of Cohen’s Kappa for each question (Cohen, 1988). According to these analyses, the two evaluators generally agreed on the response scores (κ = .88 for the first question, κ = .97 for the second question; κ = .87 for the third question, and κ = .91 overall). Disagreements were discussed and shared solution was achieved on each
Driving Hazard Perception Test in a Spanish Setting

occasion. As a high level of agreement was reached, a single researcher scored the remaining scripts (as carried out by Jackson et al. 2009).

Average scores were then calculated separately for the 20 hazardous situation clips (i.e., the hazard perception test) and for the 8 quasi-hazardous situation clips. It should be noted that the scores obtained with these two types of video clip (hazardous and quasi-hazardous situations) were intended to measure separate issues and the interpretation of the scores is actually different (respectively, the correct identification of a hazardous traffic situation and the tendency to respond that there is hazard in a traffic situation even if this is not the case). Therefore, no joint score comprising all 28 video clips was obtained.

The scores obtained from the hazardous and quasi-hazardous situation clips were submitted separately to classic item analysis and reliability analysis. In particular, a minimum acceptable item-total correlation was set at .30 and a minimum Cronbach’s alpha of .70 for each total scale was expected.

Design

The current study incorporated two separate designs. The first was a 3 x 2 mixed design, comparing performance in the hazard prediction test across 3 driver groups (learner, novice and experienced) and 2 types of hazard (environmental prediction and behavioural prediction). A second 2x2 mixed analysis was planned comparing traffic law violators to non-violators across the two types of hazard. The dependent variable in both cases was the score awarded to participants for each clip.

On the basis of their answers to 3 questions (What is the hazard? Where is it located? WHN?), each participant could score 6 points per clip.

3x2 mixed-model ANOVAs were then used to examine the differences between the groups of drivers varying in experience (learner, novice and experienced drivers) and the type of traffic situation (Environmental Prediction and Behavioural Prediction situations) for the total scores obtained from either the hazardous or the quasi-hazardous situation clips. The level of statistical significance was set at .05. The assumption of homogeneity of variance was assessed by Levene’s test. Tukey’s test for multiple comparisons was used to control overall significance. Similarly, 2x2 mixed-model
ANOVAs were used to examine the differences between the groups of drivers varying in driving recidivism (multiple offenders and non-recidivist drivers) and the type of traffic situation (Environmental Prediction and Behavioural Prediction situations) for the total scores obtained from either the hazardous or the quasi-hazardous situation clips. Statistical analyses were performed using IBM SPSS Statistics v19 for Windows.

RESULTS

Descriptive statistics and psychometrics

Descriptive statistics of the scores obtained from all the video clips are shown in Table 3 (hazardous situations) and in Table 4 (quasi-hazardous situations). Corrected item-total correlations are also shown in these tables. The item-total correlation analysis revealed that the scores from some clips were poorly associated with the corresponding total score ($r < .30$) and, consequently, they were removed from subsequent analyses. The scores from the remaining clips were averaged across participants into total scores for each scale: (a) the hazard perception test including 11 hazardous situation clips and (b) 6 quasi-hazardous situation clips. First, a higher total score from the hazardous situation clips in the hazard perception test represents a more correct evaluation of the hazardous situations (thus, higher hazard perception) and vice versa. The average score across participants was 22.88 (34.5%) (St. Dev. = 9.16). Cronbach’s alpha for the hazard perception scale was .70. Second, a higher total score with the quasi-hazardous situation clips implies a more correct assessment of the quasi-hazardous traffic situations. In this case, the average score across participants was 17.51 (48.6%) (St. Dev. = 6.99). Cronbach’s alpha for the quasi-hazardous situations was .77.

Differences in hazardous situation clips

Results in the 3 (experience: learner / novice / experienced drivers) × 2 (traffic situation: Environmental Prediction / Behavioural Prediction) mixed-model ANOVA for the score obtained from the hazardous situation video clips (see Figure 1A) allowed rejecting the null hypothesis for both
Driving Hazard Perception Test in a Spanish Setting

main effects. A significant main effect of driving experience was found \((F(2,41)=3.26; p=.049; \eta^2 p=.14)\). Tukey’s post-hoc tests suggested that the average score per clip in the hazard perception test was lower for learner drivers (1.63) than for novice drivers (2.27) and experienced drivers (2.20). In addition, statistically significant differences were found between the two traffic situations \((F(1,41)=48.08; p<.001; \eta^2 p=.54)\), suggesting that participants obtained higher hazard perception scores in Behavioural Prediction situations (2.52) than in Environmental Prediction traffic situations (1.52). The interaction term was not statistically significant in this analysis \((F(2,41)=1.79; p=.18; \eta^2 p=.08)\).

Please, Insert Figure 1 about here

The analysis in the 2 (recidivism: multiple offenders / non-recidivist drivers) × 2 (traffic situation: Environmental Prediction / Behavioural Prediction) mixed-model ANOVA for the score obtained from the hazardous situation video clips (see Figure 1B) allowed rejecting the null hypothesis for the traffic situation main effect. A statistically significant difference \((F(1,57)=76.83; p<.001; \eta^2 p=.57)\) was found between the scores in Behavioural Prediction situations (2.64) as compared to Environmental Prediction situations (1.62). The difference in the hazard perception scores between multiple offenders (2.05) and non-recidivist drivers (2.21) failed to reach the level of statistical significance \((F(1,57)=.67; p=.42; \eta^2 p=.01)\). Also, the interaction term was not statistically significant \((F(1,57)=1.20; p=.28; \eta^2 p=.02)\).

Differences in quasi-hazardous situation clips

Regarding the total scores obtained from the quasi-hazardous situation video clips, results in the 3 (experience) × 2 (traffic situation) mixed-model ANOVA (see Figure 1C) allowed rejecting the null hypothesis for the experience main effect. The overall analysis suggested statistically significant differences in the experience groups \((F(2,41)=7.78; p=.01; \eta^2 p=.19)\) and Tukey’s post-hoc analysis showed that the average score was lower for learner drivers (2.38) than for experienced drivers (3.64). There was no statistically significant difference \((F(1,41)=.002; p=.96; \eta^2 p<.001)\) between the
Driving Hazard Perception Test in a Spanish Setting

participants’ scores in Environmental Prediction (3.05) and Behavioural Prediction traffic situations (3.06). Similarly, the interaction term was not statistically significant ($F(2,41)=.32; p=.73; \eta^2 p=.02$).

Finally, results in the 2 (recidivism) × 2 (traffic situation) mixed-model ANOVA for the score obtained from the quasi-hazardous situation video clips (see Figure 1D) allowed rejecting the null hypothesis for the recidivism main effect. A statistically significant difference ($F(1,57)=5.68; p=.02; \eta^2 p=.09$) was found between the scores in the two recidivism groups, suggesting that multiple offenders were less able to correctly identify the quasi-hazardous situations (2.70) as compared to the non-recidivist drivers (3.38). In addition, there was no statistically significant difference ($F(1,57)=.05; p=.83; \eta^2 p<.001$) between the participants’ scores in Environmental Prediction (3.02) and Behavioural Prediction traffic situations (3.06). Similarly, the interaction term was not statistically significant ($F(1,57)=.13; p=.72; \eta^2 p=.002$).

DISCUSSION

The aim of the current study was to develop a hazard perception test to be applied to the Spanish driving population and to provide validity evidence for the test. To obtain evidence of the test’s validity, the potential differences between different drivers’ ability to identify hazardous and quasi-hazardous situations were explored according to drivers’ profiles of recidivism and driving experience. In particular, the results in the study suggest that some groups of drivers, such as learner drivers and multiple driving offenders, might be less able than experienced drivers or non-recidivist drivers, respectively, to correctly identify quasi-hazardous traffic situations.

To achieve this objective, a new video-based hazard perception test was developed in accordance with previous research and recommendations (e.g., Crundall et al. 2010, 2012; Jackson et al. 2009, Wetton et al. 2011). That is, individual differences in reaction time were not considered relevant when marking responses to the questions as correct; all the hazardous situation clips presented genuine driving situations. Moreover, following Jackson et al. (2009), participants had to answer different questions after each video clip to analyse the detection of the hazard as well as the driver’s situation awareness (Endsley, 1995).
Psychometric results confirm that after excluding some inconsistent video clips, the new test has adequate psychometric properties to be a useful tool for studying hazard perception in the Spanish driving population. The final version of the hazard perception test is composed of 11 hazardous situation video clips, for which an overall Cronbach’s alpha of .77 was obtained. Additionally, to obtain further validity evidence, the ability of the hazard perception test to distinguish between drivers with different experience was examined. As expected, learner drivers obtained lower average scores than novice drivers, suggesting that the ability to correctly identify hazardous traffic situations may develop early by accumulating initial driving experience. These results are mainly consistent with previous evidence obtained by Jackson et al. (2009) and Crundall et al. (2010, 2012).

In relation to the two groups of drivers varying in recidivism, results in the current study suggest that multiple driving offenders obtained similar average scores to the non-recidivist drivers in the hazard perception test when participants were shown a number of hazardous situation videos.

In addition to the hazard perception clips, participants were also shown a number of quasi-hazardous situation videos to analyse the drivers’ potential tendency to respond that there is a hazard. These clips were similar to those used in the hazard perception test, with the exception that during the recorded driving session, the potentially hazardous situation developed without involving any final hazard (i.e., the driver didn’t actually have to decelerate or make any evasive manoeuvre to avoid a potential collision). After psychometric analyses, 6 quasi-hazardous situation videos were used to compute participants’ average total scores and an overall Cronbach’s alpha of .77 was obtained. A higher total score with the quasi-hazardous situation clips implies a more correct assessment of the quasi-hazardous traffic situations. According to the analyses of the quasi-hazardous situation clips, the group of learner drivers had lower average scores than the experienced drivers.

Consequently, the overall results in the current study suggest that learner drivers are less able to correctly identify either hazardous (as previously discussed) or quasi-hazardous traffic situations. McKenna and Crick (1991) found that expert drivers were faster at detecting the hazards in videos of traffic scenes. Other authors (Finn and Bragg, 1986; McGowan and Banbury, 2004; Fisher, Pollatsek and Pradhan, 2006; Jackson et al. 2009; Underwood et al. 2013) who explored the role played by the
third level of situation awareness in Hazard Perception found significant differences between novice and experienced drivers. That is, the experienced drivers were more accurate at predicting future hazards. A potential explanation for this result might be that their accumulated experience allows more experienced drivers to better identify both kinds of situation, using their current knowledge to correctly catalogue a traffic situation as hazardous or non-hazardous. However, it is also possible that the learner drivers tested in hazard perception in the current study felt under close scrutiny and therefore tended to respond that there was a hazard even when this was not the case. This latter suggestion might imply a potential difference in their tendency to respond that there is hazard in a traffic situation even if this is not the case. One argument against the typical hazard perception methodology is that a greater bias to respond to potential hazards can increase the probability of a participant responding within an appropriate time window for an inappropriate reason. In order to overcome this problem of heightened response bias, the official UK hazard perception test reportedly disregards the responses of any clips with an excessive number of clicks. The current test, however, has demonstrated that despite the learner drivers having a greater propensity to regard the next event as a hazard (even when it was not) this did not inflate their score in the test. This supports the suggestion that ‘WHN?’ tests are less influenced by response bias than traditional response time measures.

Learner drivers had more difficulty than novices and experts in perceiving hazardous situations. This means that training in the ability to perceive hazardous situations could be given in the early stages of driving experience. Box and Wengraf (2013) argued that once a new driver has gained 1,000 miles of on-road experience, their skills and safety can be considered equivalent to those of drivers with three or more years’ experience and that experience quickly reduces the crash risk for all age groups. However, drivers require more practice to become proficient at more complex tasks (such as perceiving quasi-hazardous situations) than to reach a similar level of skill in simple tasks (perceiving some hazardous situations). The ability to perceive quasi-hazardous situations, similar to the hazardous ones, is even more difficult and requires more driving practice. These situations produce more uncertainty and recognising them is more complicated. This skill can be found later in
Driving Hazard Perception Test in a Spanish Setting

the process of acquiring driving skills, so inexperience could make the task harder (Logan, 1988). Learners and novice have less reliable memories in which to trust. Relying on memories of previously successful solutions becomes the best alternative only after years of driving experience. As a result, the experienced driver’s performance seems to be less effortful. It leads, therefore, to less disruption by distraction and more consistency (Cireri and Rucio, 2014).

Regarding the two groups of drivers varying in recidivism, results showed that multiple driving offenders obtained lower average scores in the quasi-hazardous situation videos as compared to the non-recidivist drivers. Consequently, the overall results in the current study suggest that multiple driving offenders may be less able to correctly identify quasi-hazardous situations, whereas they showed a similar performance with hazardous situation videos. A potential explanation of these results might be that multiple offenders and non-recidivist drivers varied in their tendency to respond that there was a hazard (and might not differ in their actual hazard perception per se). Although recidivist drivers appeared to be less able to correctly identify the quasi-hazardous traffic situations, it is possible that they felt under scrutiny and thus tended to respond that there was a hazard even in quasi-hazardous situations (in fact, it should be noted that the multiple offender participants were tested while attending a driving re-education course after losing their driving licence).

Both groups of drivers (multiple offenders and non-recidivist drivers) were composed of similar numbers of novice and experienced drivers and thus, the differences in driving experience might not account for the observed differences in their answers to the hazardous and quasi-hazardous situation clips. However, it should be noted that these two groups of drivers actually constitute very different populations (for example, the drivers attending re-education courses in Spain are largely men) and it is still possible that some other potential bias might have influenced the observed results. Complementary analyses were made to evaluate potential differences between the two groups of drivers recruited (see also Table 2). Results showed that the multiple offender group included more males, more people with only primary school education and fewer people with higher education. In addition, as expected, they had previously lost their driving licenses more times and had received more tickets during the previous 12 months. They also tended to drive more kilometres per year and
Driving Hazard Perception Test in a Spanish Setting

tended to be involved in more accidents during the previous 12 months (it should be noted that the multiple offender drivers had been banned from driving for some months before attending the re-education course). In consequence, the differences in these and other socio-demographic variables should be carefully considered in future studies, since they may potentially influence the ability to correctly identify hazardous and/or non-hazardous situation video clips.

CONCLUSION

The current work shows new data about how a sample of Spanish drivers behaves in the new Hazard Perception test. The test has adequate psychometric properties. In addition, we provide validity evidence for the test, based on their relation to other theoretical and practical criteria, such as the comparisons between different types of driver through different experimental conditions. The test is useful in distinguishing between learner, novice and expert drivers and it is useful in that it analyses the performance of both safe drivers (non-offenders) and unsafe drivers (re-offenders who have already lost their driving licence).

The article analyses the drivers’ performance not only when hazardous situations are shown, but also when quasi-hazardous situations are displayed. A hazardous driving situation is one where it is necessary to perform an evasive action, for instance, braking or swerving abruptly. The quasi–hazardous situations are similar to the hazardous ones, except that they do not require the driver to make any evasive manoeuvre, merely to carry on driving at the same speed and following the same trajectory.

As expected, learner drivers obtained lower average scores than novice and experienced drivers with the hazardous situation videos and learner drivers obtained lower average scores than experienced drivers with the quasi-hazardous situation videos, suggesting that the ability to correctly identify hazardous traffic situations may develop early by accumulating initial driving experience; the ability to correctly identify quasi-hazardous situations may develop later by accumulating further driving experience. In relation to the two groups of drivers varying in recidivism, results in the current study suggest that multiple offenders obtained similar average scores to non-recidivist drivers in the hazard perception situations. However, a different pattern was found with the quasi-hazardous
Driving Hazard Perception Test in a Spanish Setting

situations. Multiple offenders obtained lower average scores than non-recidivist drivers in the quasi-hazardous situation videos. As detecting traffic hazards is essential in driving, this task can be used not only as an assessment tool for novice drivers and those who do not drive safely, but also as a training tool to improve the skills of those drivers lacking the ability to detect hazards and distinguish them from quasi-hazardous situations.

ACKNOWLEDGMENTS

We wish to thank the Spanish drivers who volunteered for this study. We would like to thank the Dirección General de Tráfico (Research Project Exp: 0100DGT21263) for supporting the present research and the driving schools Victoria and Luna for their collaboration in the recruitment of participants. Financial support was provided by the Ministerio de Ciencia e Innovación, MICINN (PSI2010-15883 and PSI2011-29504) and the Junta de Andalucía (P07-SEJ-02613, P10-SEJ-6569, and Proyecto Motriz P11-SEJ-7404), FEDER funding. None of the funding sources had direct involvement in the study design, in data collection, analysis or interpretation, in the writing of reports or in the decision to submit the paper for publication.

REFERENCES


Box, E., and Wengraf, I. Young driver safety: Solutions to an age-old problem. Download from internet, on 6 September 2013. Available in: RAC
Driving Hazard Perception Test in a Spanish Setting


Driving Hazard Perception Test in a Spanish Setting


Driving Hazard Perception Test in a Spanish Setting


McKenna, F.P., and Crick, J.L. *Hazard perception in drivers: a methodology for testing and training*. Transport and Road Research Laboratory. 1991. TRL. Crowthorne, UK.


Driving Hazard Perception Test in a Spanish Setting

### Table 1. Socio-demographic information for the three experience groups of drivers (learner, novice and experienced drivers)

<table>
<thead>
<tr>
<th>Socio-demographic information</th>
<th>Learner drivers</th>
<th>Novice drivers</th>
<th>Experienced drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>14</td>
<td>18</td>
<td>36</td>
</tr>
<tr>
<td>Gender a</td>
<td>14</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Level of education b</td>
<td>14</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Years driving regularly c</td>
<td>14</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Years since obtaining driving license</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Driving frequency d</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Kilometres driven last 12 months</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Accidents with material damage last 12 months</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Accidents with victim last 12 months</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Quasi-accidents last 12 months</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Traffic incidents reported to insurance company</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Number of times losing driving license</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Traffic tickets received</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

(a) 1 = Female, 2 = Male. Median valued reported.
(b) 1 = Primary, 2 = Secondary (compulsory), 3 = Secondary (non-compulsory), 4 = Vocational, 5 = Grade, 6 = Master. Median value reported.
(c) 1 = Learning to drive, 2 = Up to 2 years, 3 = Between 3-7 years, 4 = 8 or more years. Median value reported.
(d) 1 = Daily, 2 = Weekly, 3 = Monthly, 4 = Never. Median value reported.
### Table 2. Socio-demographic information for the two recidivism groups of drivers (multiple offenders and non-recidivist drivers)

<table>
<thead>
<tr>
<th>Socio-demographic information</th>
<th>Multiple offenders</th>
<th>Non-recidivist drivers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td>Gender(^a)</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td>Level of education(^b)</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td>Years driving regularly(^c)</td>
<td>29</td>
<td>1</td>
</tr>
<tr>
<td>Years since obtaining driving license</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>Kilometres driven last 12 months</td>
<td>23</td>
<td>0</td>
</tr>
<tr>
<td>Accidents with material damage last 12 months</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>Accidents with victim last 12 months</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>Quasi-accidents last 12 months</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>Traffic incidents reported to insurance company</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>Number of times losing driving license</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>Traffic tickets received</td>
<td>27</td>
<td>1</td>
</tr>
</tbody>
</table>

\(^a\) 1 = Female, 2 = Male. Median value reported.
\(^b\) 1 = Primary, 2 = Secondary (compulsory), 3 = Secondary (non-compulsory), 4 = Vocational, 5 = Grade, 6 = Master. Median value reported.
\(^c\) 1 = Learning to drive, 2 = Up to 2 years, 3 = Between 3-7 years, 4 = 8 or more years. Median value reported.
\(^d\) 1 = Daily, 2 = Weekly, 3 = Monthly, 4 = Never. Median value reported.
Table 3. Descriptive statistics and psychometric results of the scores from the hazardous situation video clips (hazard perception test). Mean, standard deviation and corrected item-total correlation are shown for each clip score.

<table>
<thead>
<tr>
<th>Hazardous situation clips</th>
<th>Video clip</th>
<th>Mean score</th>
<th>Standard deviation</th>
<th>Corrected item-total correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>3.07</td>
<td>1.24</td>
<td>.32</td>
</tr>
<tr>
<td></td>
<td>2*</td>
<td>2.73</td>
<td>1.60</td>
<td>.21</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>1.67</td>
<td>1.12</td>
<td>.31</td>
</tr>
<tr>
<td></td>
<td>4*</td>
<td>3.36</td>
<td>1.44</td>
<td>.21</td>
</tr>
<tr>
<td></td>
<td>5*</td>
<td>1.42</td>
<td>2.08</td>
<td>.06</td>
</tr>
<tr>
<td></td>
<td>6*</td>
<td>.36</td>
<td>1.17</td>
<td>.25</td>
</tr>
<tr>
<td></td>
<td>7*</td>
<td>2.81</td>
<td>1.60</td>
<td>.26</td>
</tr>
<tr>
<td></td>
<td>8*</td>
<td>2.00</td>
<td>1.85</td>
<td>.16</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.67</td>
<td>1.29</td>
<td>.38</td>
</tr>
<tr>
<td></td>
<td>10*</td>
<td>0.68</td>
<td>0.94</td>
<td>.23</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>1.33</td>
<td>1.69</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>2.82</td>
<td>1.66</td>
<td>.36</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>2.21</td>
<td>2.40</td>
<td>.45</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>0.99</td>
<td>1.60</td>
<td>.35</td>
</tr>
<tr>
<td></td>
<td>15*</td>
<td>0.18</td>
<td>0.56</td>
<td>.00</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>3.68</td>
<td>1.27</td>
<td>.41</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>2.07</td>
<td>1.73</td>
<td>.32</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>2.97</td>
<td>2.35</td>
<td>.30</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>1.40</td>
<td>1.53</td>
<td>.33</td>
</tr>
<tr>
<td></td>
<td>20*</td>
<td>1.05</td>
<td>1.12</td>
<td>.06</td>
</tr>
</tbody>
</table>

* Hazardous situation clips excluded from subsequent analyses ($r < .30$)
Table 4. Descriptive statistics and psychometric results of the scores from the quasi-hazardous situation video clips. Mean, standard deviation and corrected item-total correlation are shown for each clip score.

<table>
<thead>
<tr>
<th>Video clip</th>
<th>Mean score</th>
<th>Standard deviation</th>
<th>Corrected item-total correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>0.07</td>
<td>0.30</td>
<td>.12</td>
</tr>
<tr>
<td>2</td>
<td>2.92</td>
<td>1.57</td>
<td>.53</td>
</tr>
<tr>
<td>3*</td>
<td>0.21</td>
<td>0.67</td>
<td>.18</td>
</tr>
<tr>
<td>4</td>
<td>2.45</td>
<td>1.44</td>
<td>.56</td>
</tr>
<tr>
<td>5</td>
<td>3.30</td>
<td>1.47</td>
<td>.53</td>
</tr>
<tr>
<td>6</td>
<td>3.48</td>
<td>1.63</td>
<td>.43</td>
</tr>
<tr>
<td>7</td>
<td>2.48</td>
<td>1.83</td>
<td>.56</td>
</tr>
<tr>
<td>8</td>
<td>2.88</td>
<td>2.15</td>
<td>.56</td>
</tr>
</tbody>
</table>

* Quasi-hazardous situation clips excluded from subsequent analyses ($r < .30$)
Figure 1. Results obtained from the hazardous and the quasi-hazardous situation video clips. Average scores across participants and standard error bars are shown for the different analyses: (A) 3 (experience) × 2 (traffic situation) with hazardous situation clips; (B) 2 (recidivism) × 2 (traffic situation) with hazardous situation clips; (C) 3 (experience) × 2 (traffic situation) with quasi-hazardous situation clips; and (D) 2 (recidivism) × 2 (traffic situation) with quasi-hazardous situation clips.
Figure 1