Transportation Research Part F 22 (2014) 218–231

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

A comparative study of two desktop hazard perception tasks suitable for mass testing in which scores are not based on response latencies

Willem P. Vlakveld *

SWOV Institute for Road Safety Research, PO Box 93113, 2509 AC The Hague, The Netherlands

ARTICLE INFO

Article history:
Received 6 March 2013
Received in revised form 12 December 2013
Accepted 13 December 2013

Keywords:
Hazard perception
Novice drivers
Driving assessment
Expertise

ABSTRACT

In PC-based hazard perception tests scores are traditionally based on how quickly participants respond to developing hazards in video clips. A disadvantage of this method is that latent hazards which do not develop into acute threats cannot be included in the test. The present study compared two tasks using the same stimuli but with different response methods. The stimuli consisted of thirteen animated video clips in which latent hazards did not materialize. Latent hazards could either be a visible other road user who due to the circumstances could start to act dangerously, or a hidden other road users who could be on collision course. The first-mentioned were the overt latent hazards and the latter were the covert latent hazards. In Task 1, participants had to indicate what the high priority latent hazard was *after* they had watched a clip. In Task 2, participants could indicate latent hazards *while* they were watching a clip and decide afterwards which of the indicated latent hazards had the highest priority. In both tasks the scores were based on how many high priority latent hazards were detected and were not based not on response times. Professional drivers (driver trainers and driving examiners) and learner drivers were randomly assigned to a group that performed Task 1 and a group that performed Task 2. Professionals scored significantly better on both tasks than learner drivers. Although in both tasks professionals scored significantly higher, Task 1 seems to be a more promising alternative for the traditional hazard perception test than Task 2 because professional drivers scored significantly higher on overt latent hazards than learner drivers in Task 1 but not in Task 2 and experience with computer games influenced the scores in Task 2 but not in Task 1. A weakness of Task 1 was its rather low internal consistency ($\alpha = .69$).

1. Introduction

Hazard Perception (HP) can be defined as situation awareness for dangerous situations in the traffic environment (Horswill & McKenna, 2004). Endsley (1995) defines situation awareness as the perception of the elements in the current situation (level 1), the comprehension of the current situation (level 2) and the projection of the future actions of the elements in the current situation (level 3). Therefore, HP can also be defined as the ability to detect (level 1) and recognize (level 2) possible hazards and to predict (level 3) how these possible hazards can develop into situations in which a crash would be very likely. It is important to note that hazard perception is about possible hazards and not about unexpected acute

* Tel.: +31 70317 3361; fax: +31 70320 1261.
E-mail address: Willem.Vlakveld@swov.nl

1369-8478/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.trf.2013.12.013

threats that require immediate action (e.g. hard braking or swerving) in order to avert a crash at the very last moment. Therefore, hazard perception is about the detection and recognition of latent hazards that are more or less hidden in the traffic scene and which can materialize, but will not always do so.

Young novice drivers are overrepresented in car crashes (e.g. OECD., 2006) and examination of crash reports has revealed that poor HP-skills play a major role in crashes involving young novice drivers (Curry, Hafetz, Kallan, Winston, & Durbin, 2011; McKnight & McKnight, 2003). Interactive PC-based training programs and training programs in which a simple driving simulator is used, enhance the skill to detect and recognize road and traffic situations which are potentially dangerous (Chapman, Underwood, & Roberts, 2002; Fisher, Pollatsek, & Pradhan, 2006; Isler, Starkey, & Williamson, 2009; Ivancic & Hesketh, 2000; McKenna & Crick, 1997; McKenna, Horswill, & Alexander, 2006; Pollatsek, Narayanaan, Pradhan, & Fisher, 2006; Pradhan, Pollatsek, Knodler, & Fisher, 2009; Regan, Triggs, & Godley, 2000; Vlakveld et al., 2011; Wang, Zhang, & Salvendy, 2010; Wetton, Hill, & Horswill, 2013). Although these training programs seem to improve one's HP-skills, their effect on the crash rate has not yet been investigated. However, an association between scores on HP-tests and crash rate has been found (Congdon, 1999; Darby, Murray, & Raeside, 2009; McKenna & Horswill, 1999; Pelz & Krupat, 1974; Quimby, Maycock, Carter, Dixon, & Wall, 1986; Wells, Tong, Sexton, Grayson, & Jones, 2008). Drivers with high scores on a HP-test have a lower crash rate than drivers with low scores. If it is assumed that HP training programs can reduce the crash rate, then it is possible that crash rates for novice drivers will decline when these training programs are incorporated into basic driver education. HP training programs will become an integral part of basic driver education when HP-skills are tested in the driving test. Testing HP-skills in the driving test is only likely to be effective if the test is valid and reliable and when it is impractical to pass the test with the aid of heuristics which have nothing or little to do with HP.

Traditionally, HP is measured with the aid of short video clips filmed from the driver's perspective. Participants who watch these clips have to imagine they are the driver of the 'camera car' (i.e. the car from which the clip is recorded). Each clip contains at least one latent hazard that partly materializes as the traffic scene evolves. Situations do not end in a collision but they develop into critical situations that require immediate action. Other road users in the scene (e.g. a pedestrian who steps into the road from between parked cars) always stop or change course just in time. Participants have to press a button as quickly as possible once they have identified a hazard. HP-skills are determined by the time that has elapsed between the first indications of a developing hazard and pressing the button (e.g. Chapman & Underwood, 1998; McKenna & Crick, 1997; Sagberg & Bjørnskau, 2006). The time that has elapsed is the Response Time (RT). In some versions of the traditional HP-test, participants do not press a button, but press one of their fingers to a touch screen at the spot of the developing hazard (Scialfa et al., 2011; Wetton et al., 2010). In this manner ambiguity is reduced about why participants press the button.

A limitation of the traditional way of testing HP-skills is its susceptibility for heuristics in doing the test which have little or nothing to do with HP. When the test is incorporated in the theory part of the driving test, the only feedback a candidate receives is if she or he has passed or failed the entire test. Yet indirectly a candidate may get feedback per hazard in each video clip. A candidate may miss a latent hazard in its early stage of development. This candidate will learn what she or he has missed because the latent hazard develops into an acute threat. Threatening driver situations are better remembered than non-threatening driving situations (Groeger, 2000). This also holds for threating situations experienced in a simulator (Koustanaï, Boloix, Van Elslande, & Bastien, 2008). If a candidate passes on this information to candidates who have not yet done the test, these new candidates will have an advantage when it is their turn to do the test. In the present study response methods for HP-tasks were investigated which are less likely to be susceptible to fraudulent heuristics because the latent hazards do not develop into acute threats.

Besides its susceptibility in passing the test without HP-skills, the traditional way of testing HP-skills has other weakness. Its criterion validity is not undisputed and the traditional HP-test cannot discriminate between various types of HP-skills. Although in most studies the average RT appeared to be longer for novice drivers than for experienced drivers (McKenna & Crick, 1997; Scialfa et al., 2011; Smith, Horswill, Chambers, & Wetton, 2009; Wetton, Hill, & Horswill, 2011; Wetton et al., 2010), there are also studies in which no difference in average RT was found (Chapman & Underwood, 1998; Crundall, Underwood, & Chapman, 2002; Sagberg & Bjørnskau, 2006). Reaction times increase as one grows older (e.g. Deary & Der, 2005). Experienced drivers are normally older than novice drivers and will therefore have longer reaction times. However, older, more experienced drivers may still have shorter RTs than young novice drivers because young novice drivers have not yet developed elaborated schemata (mental representations). Elaborated schemata enable drivers to detect and recognize latent hazards as the road and traffic scene unfolds (Norman & Shallice, 1986; Vlakveld, 2011). However, it is possible that the shorter RTs of experienced drivers are not caused by better HP-skills but are caused by differences in risk acceptance. That is, older drivers may press earlier because they are more cautious than younger drivers. If it was not due to the not yet fully elaborated schemata, it could be that in some studies no effect was found because the latent hazards in the tests were so obvious that novice drivers could also easily recognize them. Or, on the other hand, it could also be the result of latent hazards in these tests being so unexpected that experienced drivers could not predict them either. If the cause is a difference in risk acceptance between the two groups, novice drivers would also have had longer RTs when hazards are very obvious. Although Sagberg and Bjørnskau (2006) did not find an overall difference in RTs between novice drivers and experienced drivers, they found shorter RTs for experienced drivers in three situations with more complex latent hazards. This is an indication that better developed schemata are a more likely explanation for shorter RTs of experienced drivers than differences in risk acceptance. Another indication that the longer RTs of novice drivers are probably caused by poor HP-skills, is that Wallis and Horswill (2007) did not find differences in risk ratings of hazard scenes between novice and experience drivers whereas they did find that experienced drivers had shorter RTs than novice drivers.