Accident Analysis and Prevention 51 (2013) 268–273

Contents lists available at SciVerse ScienceDirect

Accident Analysis and Prevention

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

A comparison of static and dynamic hazard perception tests

Charles T. Scialfa*, David Borkenhagen, John Lyon, Micheline Deschênes

Department of Psychology, University of Calgary, Calgary, Alberta T2N 1N4, Canada

ARTICLE INFO

Article history:
Received 9 October 2012
Received in revised form 3 December 2012
Accepted 4 December 2012

Keywords:
Hazard perception tests
Driver safety

ABSTRACT

Hazard perception is a vital component to safe driving and hazard perception tests (HPTs) are being used with greater frequency for driver training, assessment and licensure. In this study, we compared a dynamic HPT (Scialfa et al., 2011), which presents short video scenes to observers and a static HPT (Scialfa et al., 2012), which uses still images. Both tests require the observer to indicate the presence of a traffic conflict that would lead to a collision between the "camera" vehicle and another road user or fixed object. Young adult drivers ($n = 56$) completed both forms of the HPT, along with a modified version of the Driver Behavior Questionnaire (Reason et al., 1990) and a measure of simple reaction time. Self-reported collision and moving violation data were also collected. As in previous work, both static and dynamic HPTs had good reliability. The correlation between composite static and dynamic scores was approximately .40, but was reduced to approximately .25 when simple reaction time was controlled. Both HPTs predicted lapses and errors on the Driver Behavior Questionnaire, but neither predicted self-reported collisions or moving violations. Discussion focuses on the differences in visual cues available in dynamic and static scenes and how these differences could influence decisions about potential hazards.

Among the many skills required for safe driving, the accurate and efficient perception of hazards is certainly one of the most critical. As such, the expanding research literature on hazard perception tests (HPTs) for licensing, assessment and training has garnered attention from researchers and policy-makers alike. HPTs have been developed in a variety of formats, including those utilizing still images (Huestegge et al., 2010; Scialfa et al., 2012; Whelan et al., 2002) or simulated plan views of potentially hazardous scenarios (Fischer et al., 2006), dynamic video sequences (Horswill et al., 2010b; Scialfa et al., 2011; Shahar et al., 2010) and simulations (Pollatsek et al., 2006). The increased use of HPTs in these various forms has prompted questions about their congruence, relative advantages and disadvantages. In the present study, we compared two HPTs, one involving still images and another using dynamic sequences. In doing so, we were able to determine the reliability of each test, as well as the zero-order correlation between these two measures of hazard perception. We were also able to estimate the association between these two forms after eliminating variance due to simple reaction time. Finally, ancillary measures using the Driver Behavior Questionnaire (Reason et al., 1990) and self-reported driving incidents allowed us to compare a dynamic HPT (DHPT) and static HPT (SHPT) in criterion-based validity.

Hazard perception, defined as drivers' ability to anticipate dangerous situations on the road ahead (Horswill and McKenna, 2004), has been related causally to collision risk (Insurance Institute for Highway Safety, 2010; McKnight and McKnight, 2003) and performance on HPTs predicts crash risk in diverse populations (Darby et al., 2009; Horswill et al., 2010a; Wells et al., 2008). As well, two groups with elevated crash risk, novice and older drivers, are less able to identify and respond to hazards than experienced younger or middle-age drivers (Horswill et al., 2009; McKenna and Crick, 1994; Pollatsek et al., 2006; Quimby and Watts, 1981; Wetton et al., 2010; Whelan et al., 2002; Scialfa et al., 2011, 2012; but see Kaber et al., 2012). Results such as these have facilitated the implementation of HPTs as part of the licensure process (see, for example, Wells et al., 2008).

Many of the studies cited above have used dynamic video sequences or simulated scenarios to assess hazard perception and certainly, on face validity arguments, this format has obvious appeal. However, some investigators have suggested that HPTs consisting of still images may offer an effective assessment or training alternative (e.g., Huestegge et al., 2010; Whelan et al., 2002). There are several advantages to using still images. In still images it is often easier to select the moment of onset of one unambiguous hazard, whereas an extended dynamic scene may contain multiple hazards, which can create problematic response variability. Because hazard onset is synchronous with image onset in a static image, there is no uncertainty about determining the onset of the hazard, a factor

* Corresponding author.
E-mail address: scialfa@ucalgary.ca (C.T. Scialfa).

0001-4575/$ – see front matter
http://dx.doi.org/10.1016/j.aap.2012.12.006

that can complicate the calculation of response latency in dynamic scenes. In addition, a test containing still images can be completed in substantially less time or, alternatively, include a larger number of stimuli, because a typical dynamic scene lasts 30–60 s whereas a still image can have a fixed and much shorter display duration.

Another advantage of SHPTs is the ability to add or subtract elements from the images using photo-editing software. de Craen et al. (2008) created an "Adaptation Test" which consisted of 18 traffic scenes presented in two almost identical photographs which differed in one single detail, thereby increasing the complexity of the situation. Respondents were asked to assess the speed they would drive in the depicted situation. Novice drivers did not adapt their responses to more complex scenes to the same degree as more experienced drivers. Marrington et al. (2008), Wetton et al. (2010) and Velichkovsky et al. (2002) (Exp. 1) have also used still images to create a static Hazard Change Detection Task, which would also be more difficult in a dynamic format.

On the other hand, there are obvious differences between the hazards presented in a static image and those one encounters on the roadway. While visual search generally is required to identify hazards in both static and dynamic forms of the HPT, search demands are much greater when the scene contents and the interplay among scene objects is changing continually, as they do while one is driving. It has been argued that some individual differences in hazard perception arise because of differences in visual search (Borowsky et al., 2010; Falkmer and Gregersen, 2005; Pradhan et al., 2005) and so a dynamic HPT that engages search mechanisms in a manner similar to driving will be more likely to reflect these differences in search behavior.

Additionally, one can argue that the "anticipatory cues" (see Wetton et al., 2011), which are important for accurate hazard identification, are reduced or missing entirely in static scenes. In particular, motion cues that provide information about the trajectory of the driver and other road users are absent in still images. As Wetton et al. (2011) point out, observers may, therefore, find it difficult to identify the hazard or evaluate the urgency of a response. This could be particularly true for experienced drivers who are accustomed to using motion information to assess hazard potential when they are driving.

Yet another perspective suggests that responses to static images may not mirror those obtained from dynamic scenes. Underwood (2007) has invoked Endsley's (1995) situation awareness theory to describe a mental model of the roadway environment that includes perception, integration of isolated perceptions into an awareness of current driver actions and positions, and finally, prediction of future behaviors, speeds, trajectories, etc. It may be more difficult to evaluate a static scene against this mental model and, in consequence, static HPTs may be a relatively weak measure of hazard perception.

The foregoing issues motivated a comparison of static and dynamic HPTs. Two of our recent studies (Scialfa et al., 2011, 2012) developed an SHPT and a DHPT that were found to have good internal consistency as well as the ability to discriminate novice and experienced, like-aged drivers. In the current study, we gave these two HPTs to a group of young drivers to ascertain the strength of association between them. On the view that they measure similar aspects of behavior, one might expect the correlation between them to be large. However, because motion-based information is missing in the SHPT, the correlation between the tests may be quite low, particularly after controlling for response speed, which can be viewed as distinct from hazard perception as such (Wetton et al., 2011). Additionally, in order to compare the two HPT formats along the dimension of criterion-based validity, we examined their association to self-reported driving behavior, collisions and traffic violations.

1. Method

1.1. Participants

Fifty-six drivers participated in this study. All participants were 18–25 years of age, held a valid driver license or a learner's permit, were community-dwelling with good self-reported health and had normal or corrected-to-normal vision. The average age was 20.7 years (*SD* = 1.67 years). On average they were licensed for 4.5 years (*SD* = 2.12) and drove on average 10,928 km/year (*SD* = 9913). Thirteen people reported having at least one collision and seventeen people reported having at least one moving violation in the past 2 years. Participants were recruited from undergraduate courses at the University of Calgary and advertisements posted around the campus. They had the option to receive $20 (CDN) or, if they were a University of Calgary psychology student, partial course credit in exchange for their involvement. The Conjoint Faculties Research Ethics Board of the University of Calgary approved this research.

1.2. Materials and apparatus

The footage used to produce the still images for the SHPT and the video clips for the DHPT was gathered in Vancouver, BC and surrounding areas using a Sony Handycam Camcorder, model HDR-SR11. The camera was mounted inside a 2005 Subaru Impreza and secured to the inside door window on the passenger side of the vehicle. An extendable arm allowed the videotaped scenes to give an approximate "driver's eye" view. Filming occurred during daylight hours, under fair skies and dry roadway conditions. The footage was recorded in the AVCHD 16M (FH) format, which had a resolution of 1920 × 1080/60i. Still images were produced by rendering a video frame into 4 s movie clips at a display resolution of 1280 × 1024. The location and spatial coordinates of the traffic conflicts at onset were defined using 'Media' software developed at the University of Queensland (Jackson, 2008).

A 17-in. Elo TouchSystems (1729L) touch-sensitive LCD desktop monitor with a resolution of 1280 × 1024 was used to display the stimuli and collect responses. Touching the screen had no effect on the amount of time the images were shown and there was no limit on the number of times a participant was allowed to touch the screen or the number of objects a participant was allowed to touch. Reaction times to the defined traffic conflicts were defined from the first time an observer touched them.

1.3. Static hazard perception test (SHPT)

One hundred and twenty unique images of driving scenes were presented to the participants in 2, 5-minute blocks consisting of 60 images each. One hundred (83%) of the images selected were of potential traffic conflicts. Potential traffic conflicts were defined as situations in which a collision (or near collision) between the driver and another road user would occur, or had the potential to occur unless the driver took an evasive action, such as slowing, stopping, swerving or steering (see Table 1). Twenty images (17%) did not contain a potential traffic conflict (hereafter designated as No Traffic Conflict (NTC) scenes) and were included in the series to modulate the observer's criterion for making a response. Each image was presented for a duration of 4 s, followed by a 1 s delay preceding the next trial. More details can be found in Scialfa et al. (2012), where it was demonstrated to have good internal consistency and the ability to differentiate novice and experienced younger drivers.