Accident Analysis and Prevention 148 (2020) 105785

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

A cross-cultural comparison of visual search strategies and response times in road hazard perception testing

Leandro L. Di Stasi[a,1,*], Carolina Diaz-Piedra[a,b,*,1], José M. Morales[a,c], Anton Kurapov[d], Mariaelena Tagliabue[e], Anna Bjärtå[f], Alberto Megias[g], Jens Bernhardsson[f], Svitlana Paschenko[d], Samuel Romero[c], Antonio Cándido[a], Andrés Catena[a]

[a] Mind, Brain, and Behavior Research Center, University of Granada, Granada, Spain
[b] College of Nursing and Health Innovation, Arizona State University, Phoenix, AZ, USA
[c] Department of Computer Architecture and Technology, University of Granada, Granada, Spain
[d] Faculty of Psychology, Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
[e] Department of General Psychology, University of Padova, Padova, Italy
[f] Department of Psychology and Social Work, Mid Sweden University, Östersund, Sweden
[g] Department of Basic Psychology, Faculty of Psychology, University of Malaga, Malaga, Spain

ARTICLE INFO

Keywords:
Gaze entropy
Hazard perception
Regional differences
Road risk perception
Traffic safety

ABSTRACT

Road hazard perception is considered the most prominent higher-order cognitive skill related to traffic-accident involvement. Regional cultures and social rules that govern acceptable behavior may influence drivers' interpretation of a traffic situation and, consequently, the correct identification of potentially hazardous situations. Here, we aimed to compare hazard perception skills among four European countries that differ in their traffic culture, policies to reduce traffic risks, and fatal crashes: Ukraine, Italy, Spain, and Sweden. We developed a static hazard perception test in which driving scenes with different levels of braking affordance were presented while drivers' gaze was recorded. The test required drivers to indicate the action they would undertake: to brake vs. to keep driving. We assessed 218 young adult drivers. Multilevel models revealed that the scenes' levels of braking affordance (i.e., road hazard) modulated drivers' behavior. As the levels of braking affordance increased, drivers' responses became faster and their gaze entropy decreased (i.e., visual search strategy became less erratic). The country of origin influenced these effects. Ukrainian drivers were the fastest and Swedish drivers were the slowest to respond. For all countries, the decrement in response times was less marked in the case of experienced drivers. Also, Spanish drivers showed the most structured (least erratic) visual search strategy, whereas the Italians had the most rigid (most constant) one. These results suggest that road hazard perception can be defined cross-culturally, with cultural factors (e.g., traffic climate, legislation) modulating response times and visual search strategies. Our results also support the idea that a multimodal assessment methodology is possible for mass testing of road hazard perception and its outcomes would be relevant to understand how different traffic cultures shape driving behavior.

1. Introduction

Despite the fact that traffic legislations have made roads safer globally (World Health Organization, 2018), the advances reducing the number of traffic fatalities and injuries have slowed down over the last years. For example, from 2017 to 2018, the traffic death rate increased by 28% in Sweden, whereas a decrease of only 1.3% has been shown in Spain (Adminaité-Fodor et al., 2019). One of the reasons behind the disparities in the success of traffic legislations among countries would arise at the level of local attitudes and social rules towards what constitutes a safe driving behavior (Cassarino and Murphy, 2018; Meesmann et al., 2018). That is, a regional/national culture would influence what is considered normal or acceptable behavior and, consequently, drivers' interpretation of a traffic situation. For instance, Europe is the continent with the lowest death rate from road traffic accidents (9.3 per 100,000 population; World Health Organization, 2018). However, this

* Corresponding authors at: Mind, Brain, and Behavior Research Center – CIMCYC, University of Granada, Campus de Cartuja s/n, 18071, Granada, Spain.
E-mail addresses: distasi@ugr.es (L.L. Di Stasi), dipie@ugr.es (C. Diaz-Piedra).

[1] These authors should be considered co-first authors.

https://doi.org/10.1016/j.aap.2020.105785
Received 11 May 2020; Received in revised form 25 August 2020; Accepted 11 September 2020
Available online 5 November 2020

number differs greatly between the European countries. The latest data on traffic deaths across the European Union (EU-27) show that there is a widening gap between the actual and the desired progress towards the EU 2020 target ("Vision Zero"), with only one country with less than 30 deaths per million of inhabitants (Sweden), nine countries with 30–39 deaths (e.g., Spain), twelve countries with 40–67 deaths (e.g., Italy), and five countries with more than 68 deaths (Carson et al., 2020). Despite fairly consistent legislations across the Member States of the EU and significant improvements in infrastructure, there are major differences in road traffic fatalities. Therefore, understanding the different cultural thresholds for what constitutes a road hazard could be one of the main challenges to envision a successful common road traffic legislation.

1.1. Cross-cultural approaches to study drivers' behavior

Because of the intrinsic complexity and costs of investigating cross-cultural differences in road safety issues, studies have largely relied upon self-reported data (e.g., Scott-Parker and Oviedo-Trespalacios, 2017). In cross-cultural studies, large amounts of road safety data have been collected (mostly through online surveys), often involving the frequency of self-reported safe/unsafe behaviors over a passed time period. For example, de Winter and Dodou (2016) collected data from more than 6,000 respondents, comprising drivers from 41 countries. They found that self-reported non-speeding violations (for example, aggressive behaviors, or tailgating) correlated positively with national road traffic death rates. More recently, the E-Survey of Road Users' Attitudes (ESRA) project surveyed almost 40,000 road users from 38 countries (Meesmann et al., 2018). In Europe, an *a priori* homogenous region, the ESRA project found that attitudes towards speeding and distracted driving largely differed among countries (n = 17,000, 17 countries, Torfs et al., 2016). These international surveys are good examples of how traffic researchers can highlight cross-cultural differences in an efficient way. However, surveys are also well known for their dependence on personal and motivational factors that could increase the social desirability bias and inadvertently jeopardize the obtained results (Lajunen and Summala, 2003; Sullman and Taylor, 2010). Moreover, although self-reports are often used, the results about their consistency with actual driving behavior (e.g., making a timely response when a road hazard is present) are not clear (Āບele et al., 2018).

A valid, and more objective, alternative to study cross-cultural differences in safe/unsafe driving behaviors is represented by computerized *hazard perception tests*, where road hazard perception is the ability to perceive, to anticipate, and to respond to traffic situations that have a high probability to lead to a crash (e.g., Ehsani et al., 2020). Road hazard perception is not only a critical component of safe driving, but it is also considered the most prominent higher-order cognitive skill related to traffic-accident involvement (Horswill and McKenna, 2004; for a recent review, see Moran et al., 2019). Road hazard perception tests present driving scenes (dynamic [video clip] or static [picture]) filmed/taken from the driver's perspective, with a number of more or less hazardous situations that require a more or less immediate response (e.g., Pelz and Krupat, 1974; for a recent review, see Moran et al., 2019). Participants (mostly drivers) are asked to press a response button as quickly as possible once they identify a hazard or, depending on the specific aim of the test, once they have made a behavioral decision (such as to brake or not). In order to reach a behavioral decision that is appropriate, a driver first has to detect a potential hazard and then decide if it is a true hazard (Huang and Winston, 2011). This chain of decision-making events usually takes few seconds (∼ 2) (Olson and Sivak, 1986; Summala, 2000). Previous studies have shown that safer drivers tend to respond earlier and more efficiently to hazardous situations (Moran et al., 2019). Moreover, response times to hazards would improve with driving experience (Grayson and Sexton, 2002) and would be inversely associated with crash involvement (Horswill et al., 2010).

It is reasonable to assume that cultural differences, including both legal and social rules that govern acceptable behavior, may modulate the ability to detect hazards (Bazilinskyy et al., 2020; Nævestad and Bjørnskau, 2012). Sivak and colleagues were the first to show that a driver's nationality could influence both risk- perception and risk-behavior using a hazard perception test (Sivak et al., 1989a, 1989c, 1989b). In particular, participants from Spain, United States, West Germany, and Brazil differed when estimating the hazard involved in different traffic scenes. Spanish drivers estimated the highest risk, and US drivers the lowest. Such results suggest that culture may influence a driver's skills to determine the threshold of what could represent a road hazard. Since these pioneering studies, a handful of other studies have confirmed differences across countries in drivers' behavior also using alternative measures of hazard perception (for a recent study, see Lee et al., 2020). Despite the advantages of using an objective test, such as a hazard perception test, to study a driver's skills, a comparison of results obtained in different countries is difficult. An important reason is that it is rare to find studies that have employed the same methodology and sets of stimulus-material in all participating sites. One of the few studies that have applied the same methodology compared drivers from UK, Spain, and China (Ventsislavova et al., 2019). The authors found that drivers' criterion level for responding to hazards (i.e., percentage of hazards identified, response times) would be culturally sensitive, whereas the ability to predict hazards (i.e., percentage of correct predictions) would not.

All road hazard perception tests collect drivers' behavioral responses, usually response times to identify a hazard or to decide which behavior to take. Nevertheless, some authors argue that response time is a simplified description of a driver's hazard perception ability (Horswill et al., 2020) and alternative measures are needed. Visual behavior is the predominant source of information used while driving (Crundall and Underwood, 2011; Sivak, 1996). Previous studies show that encountering a hazardous event on the road affects drivers' visual behavior, eliciting a general reduction of gaze dispersion (Savage et al., 2020). As explained by Underwood and colleagues, "when a hazardous area of the scene has been potentially identified, it is important that information in this region is processed in depth. There may be advantages to monitoring that location for further unfolding of events" (Underwood et al., 2011). However, cross-cultural studies have rarely focused on this measure. In one preliminary work, authors conducted the same multimodal assessment over participating sites, comparing visual search strategies using eye tracking along with hazard identification responses, in Malaysian and UK drivers (Lim et al., 2013). Malaysian drivers had a higher threshold for what constitutes a hazard, as shown by a lower response rate to road hazards. However, differences in visual search strategies between Malaysian and UK drivers were not evident. Unfortunately, due to technological difficulties (complex calibration procedures, bulkiness of the device) along with the complexity of carrying out a cross-cultural study, the eye tracking-based methodology has not been widely adopted in mass testing since then. However, modern, affordable, and easy-to-use eye trackers have overcome many of these barriers, offering new opportunities for road safety researchers.

1.2. Our road hazard perception testing approach

Aiming to investigate cultural influences (herein roughly represented as country of origin) on driver hazard perception, we designed an international comparative study, implementing a typical hazard perception response time paradigm. Participants from different countries were assessed with the same experimental set up: procedure, stimulus material, apparatus, and experimenter. Participants assessed 140 still pictures of natural driving scenes from different European roads with different levels of road hazard. To understand the cultural impact on risk perception, both hazard identification responses and drivers' visual behavior are important. For each picture, we therefore recorded (1) the participant's option of action (braking vs. keep driving)– interpreting the braking as intended collision avoidance by speed reduction (as defined by Reinisch et al., 2010)–, (2) response time to action (hereinafter abbreviated as RT), (3) perceived levels of road hazard, and (4) visual search strategy. In order to investigate visual search strategies, we used an eye tracker to measure gaze entropy. Gaze entropy describes the amount of disorder (randomness) present in the visual search, coinciding with the concept of entropy in thermodynamics (Harris et al., 1986). Thus, we attempted to quantify the driver's scanning pattern to explain the orderliness (or lack thereof) of his/her visual search strategy (for a recent review on this measure, see Shiferaw et al., 2019).

1.3. Selected countries and working hypothesis

We focused on drivers from four European countries: Ukraine, Italy, Spain, and Sweden. These countries clearly differ in (1) their risk culture (de Winter and Dodou, 2016), (2) the success of their legislations in reducing road traffic risks (Antov et al., 2012), and, consequently, (3) in their estimated number of fatal crashes (World Health Organization, 2018). Table 1 summarizes some differences in socioeconomic and road safety data between the four countries.

Italy, Spain, and Sweden are EU Member States, but their rate of deaths from road traffic injuries differ considerably. Sweden is the country with the lowest rate in the EU. Italy and Spain, both Mediterranean countries, are, respectively, above and below the average EU road traffic fatality rate (EU 27 average 2019: 51; Carson et al., 2020). The inclusion of Ukraine, a non-EU member within the European neighborhood, is of great relevance to our study. This post-soviet country has one of the highest rates of people killed on European roads (World Health Organization, 2018). Yet, only a few studies have investigated road safety issues within this country (see recent studies by Sullman and colleagues; e.g., Hill et al., 2019; Sullman et al., 2018). Thus, comparative road safety studies that include Ukrainian drivers are relevant considering the rapid growth of flow of people (Eurostat, 2018) and trade across Ukraine and EU (Directorate-General for Trade European Commission, 2020).

Table 1
Summary of differences among Ukraine, Italy, Spain, and Sweden in socioeconomic and road safety data, number of vehicles and vehicle standards, and the existence of good laws (as defined by the World Health Organization [WHO]) that are more or less enforced (in a 0 to 10 scale) about speed, alcohol consumption, and seat-belt use. Adapted from the *Global status report on road safety 2018* (WHO, 2018).

	Ukraine	Italy	Spain	Sweden
Population (millions)	44.4	59.4	46.3	9.4
Income group[†]	Lower-middle	High	High	High
GNI per capita (US$, thousands)	2.3	31.6	27.5	54.6
National road safety strategy	No	Yes	Yes	Yes
Deaths/100k[*]	13.7	5.6	4.1	2.8
Number of vehicles (millions)	14.4	52.6	32.9	6.10
Number of key vehicle standards (out of 8)	0	8	8	8
Electronic stability control	No	Yes	Yes	Yes
Frontal impact	No	Yes	Yes	Yes
Pedestrian protection	No	Yes	Yes	Yes
Motorcycle anti-lock braking system	No	Yes	Yes	Yes
Good drink driving laws/ enforcement	Yes/5	Yes/7	No/7	Yes/6
BAC limit, general (g/dl)	≤0.02	≤0.05	≤0.05	≤0.02
BAC limit, young/novice (g/dl)	≤0.02	0.00	≤0.03	≤0.02
Good speed limit laws/enforcement	No/3	Yes/8	Yes/8	Yes/8
Speed limit, urban roads (km/h)	60	50	50	50
Speed limits, rural main roads (km/h)	90	110	90	110
Speed limits, motorways (km/h)	130	130	120	120
Good seat-belt laws/enforcement	Yes/3	Yes/7	Yes/8	Yes/6
Required in front and rear seats	Yes	Yes	Yes	Yes

Note. BAC = blood alcohol concentration; GNI = gross national income.
[*] Estimated by WHO, year 2016.
[†] As defined by the World Bank.

In accordance with the idea suggested by Lim and colleagues (2013), actual experience with more hazardous driving events may play out in two directions: drivers would become quicker and more accurate when detecting road hazards (with a more effective visual search strategy) due to a greater exposure, or alternatively, would become slower and less accurate when detecting road hazards (with a less effective visual search strategy) due to desensitization. Thus, based on the different risk cultures (de Winter and Dodou, 2016) and road safety statistics (e.g., Adminaité-Fodor et al., 2019; World Health Organization, 2018), we expected that Swedish drivers, having one of the lowest traffic death rates in the world, and Ukrainian drivers would be positioned at the opposite ends of the road hazard perception continuum. Ukrainian drivers would therefore exhibit not only the lowest perceived levels of road hazard (as shown by the number of brakes), but also the longest RTs and a higher amount of gaze dispersion, as they would explore the scenes in a less structured way. Due to similarities in road safety statistics, Italian, and Spanish drivers were expected to fall somewhere in between Ukrainian and Swedish drivers.

2. Materials and methods

2.1. Experimental design

We analyzed the promptness with which participants decided to brake or to keep driving (RTs) and the visual search strategies (i.e., gaze entropy) depending on two independent variables: (1) the drivers' *country* and (2) the average *levels of road hazard* (braking affordance, see section 2.6) present in different natural driving scenes. The effects of drivers' *gender* and *driving experience*, as well as the *visual salience* of the images were also explored. Additionally, we examined the differences among *countries* in the average frequency of the braking behavior for the whole set of pictures.

2.2. Participants

We used a convenience sampling method to recruit 236 drivers from four countries (50 Ukrainian [25 female], 56 Italian [27 female], 69 Spanish [41 female], and 61 Swedish [31 female] drivers). The incidence of road traffic injuries and fatalities is especially high among young adults (James et al., 2020). In our study, we therefore focused on active young adult drivers (the age range was 18–38 years). No power calculations were undertaken due to the lack of applicable pilot data. The number of participants was considered appropriate based on a previous cohort where statistically significant differences in oculomotor and behavioral metrics were found between Malaysian and UK drivers (Lim et al., 2013). Inclusion criteria were the following: (1) normal or corrected to normal vision, (2) having a valid domestic driving license, and (3) appropriate levels of arousal before the experimental session, operationalized as a score lower than 3 on the Stanford Sleepiness Scale (SSS, Connor et al., 2002; Diaz-Piedra et al., 2019; Hoddes et al., 1972), which indicates no fatigue and/or sleepiness. Exclusion criteria were the following: the consumption of (1) illegal habit-forming drugs (including medical cannabis) or (2) more than 40 g/day of alcoholic beverages. We excluded four participants for drugs consumption (2 Spanish, 1 Italian, 1 Swedish), three participants for having their driving licenses withdrawn (1 Italian, 2 Swedish), and another one (Swedish) after the arousal assessment.

2.3. Stimuli and experimental task

We randomly presented 140 still pictures, selected from a large and detailed image database depicting natural driving scenes ($n = 347$) (Megías et al., 2015). All pictures were taken from the driver's point of view on South and North European roads (see Supplementary material for a description of the selected scenes). As explained by Megías and colleagues, "all images met certain statistical criteria aimed to reduce interpersonal variability in the interpretation of the traffic situation and the estimated speed at which a vehicle is traveling in static traffic scenes". That is, forty driving instructors assessed the image database and the selection of the final set of pictures was based on that assessment. The set of pictures has been proven to be a valid tool to study road hazard perception by eliciting specific drivers' behavioral responses (Maldonado et al., 2020; Torres et al., 2017), electrophysiological brain patterns (Baltruschat et al., 2020; Megías et al., 2018), and neural mechanisms (Megías et al., 2015).

Each experimental trial consisted of two stages. An instruction screen (written in the corresponding language) indicated the type of task to be performed and preceded each stage. First, participants were instructed to watch the pictures and to left- or right-click on a mouse (with the left or right hand) as soon as they had decided whether to brake or to keep driving (Fig. 1, *Stage 1*). Braking decision was understood as intended collision avoidance by speed reduction (Reinisch et al., 2010). Based on the canonical brake RT (Summala, 2000), we set a maximum RT of two seconds. At the second stage, following each picture, participants assessed the perceived hazard level of the driving scene using a 7-point Likert rating scale from 0 (not hazardous) to 6 (extremely hazardous) (Fig. 1, *Stage 2*). For Stage 2, we set a maximum assessment time of five seconds (Díaz-Román et al., 2015).

We set a criterion to exclude participants who did not perform the task correctly. Eight participants, who decided to keep driving in more than 25% of the driving scenes with high levels of road hazard (average levels of perceived road hazard ≈ 4), were excluded (1 Italian, 4 Ukrainian, 3 Swedish).

2.4. Gaze recording and analyses

We recorded binocular gaze data at 30 Hz using the EyeTribe eye tracker (Oculus, Menlo Park, CA, USA). This remote eye tracker is an easy to transport and set up low-cost device that provides a 0.5° of accuracy and a spatial resolution of 0.1°. Data obtained with this device are comparable to those obtained with gold standard devices (Titz et al., 2018).

We used a custom code and the MATLAB Psychophysics Toolbox (Brainard, 1997) to display the trials, and to collect eye movement data and behavioral responses. We identified blink periods as portions of the raw data where eye information was missing for 100 ms or more. We removed these segments from the analysis as well as the 200 ms before and after each blink or semi-blink in order to eliminate the initial and final parts during which the pupils were partially occluded. As a measure of drivers' visual search strategies, we calculated gaze entropy using the classical Shannon's entropy formula (Shannon, 1948), defined as:

$$H_g(X) = -\sum p(x,y) \cdot log_2 p(x,y)$$

where H_g is the entropy value set for X (a picture for each participant) and $p(x,y)$ is the probability of the subject's gaze falling in the (x,y) position of the visual field for a given sample, estimated from the recording of each picture, measured in bits. This gives a measure of the average uncertainty over the position of the gaze on an instant in time during the exploration of the picture, or equivalently, the information provided by a single observation. In order to calculate the gaze entropy, we divided the visual field of the stimulus (~36° horizontal and ~23° vertical degrees of visual angle) in 828 bins of 1° × 1°, and calculated the probabilities of the gaze falling in each of these bins at any given time.

Due to log system failures during the recording, we discarded 1 Spanish and 1 Swedish driver. Therefore, we finally analyzed data from 218 participants out of the initial pool of 236 drivers: 66 Spanish, 53 Italian, 46 Ukrainian, and 53 Swedish drivers. Table 2 presents the characteristics of the final pool of participants for each country.

2.5. Procedure

We conducted the study in conformity with the Code of Ethics of the World Medical Association (WMA, Declaration of Helsinki) (Williams, 2008). The experiment was carried out under the guidelines of the University of Granada's Institutional Review Board (IRB approval #484/CEIH/2018).

All experimental materials were originally written in Spanish. We

Table 2
Final sample characteristics. Participants' gender, age, and driving experience (years since the initial issue of the driving license) (n = 218) for each country. Experienced drivers are those with more than 5 years of experience.

	Ukraine *n* = 46	Italy *n* = 53	Spain *n* = 66	Sweden *n* = 53
Gender, female *N*	24	27	40	29
Age, years *M* (SD)	23.91 (5.78)	23.15 (3.32)	24.09 (3.58)	28.69 (4.92)
Driving experience, years *M* (SD)	4.45 (4.67)	4.61 (3.05)	5.16 (3.64)	9.44 (5.17)
Experienced drivers, *N*	13	17	31	43

Note. *M* = mean; SD = standard deviation.

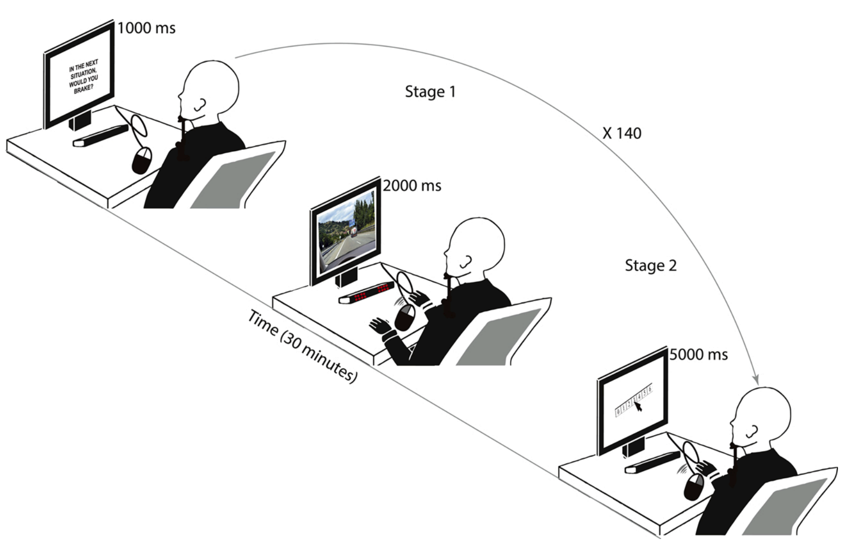

Fig. 1. Schematic representation of the experimental timeline (image, in part, adapted from Yamada and Kobayashi, 2018). Participants underwent 140 experimental trials while resting their head on a chin rest approximately 60 cm from the screen. The eye tracker was positioned below the screen. Each trial consisted of two stages with the following structure. Stage 1: (1) an instructions screen (1,000 ms); (2) a central fixation cross (750 ms); (3) a driving scene picture (2,000 ms), participants had to press a mouse button as quickly and accurately as possible to indicate whether they would have decided to brake or to keep driving given each driving scenario. Stage 2: (1) an instructions screen (1,000 ms); (2) a 7-point Likert rating scale from 0 (not hazardous) to 6 (extremely hazardous) (5,000 ms), participants had to assess the perceived hazard level of the driving scene.

did forward-backward translations to English and, then, to Ukrainian, Italian, and Swedish. Four of the researchers (native speakers) ensured the accuracy of the translations (Spanish-English: CDP, AC; English-Ukrainian: AK, SP; English-Italian: LLDS, MT; English-Swedish: AB, JB).

After the participants signed the informed consent form, they filled in several questionnaires (sociodemographic, health, and driving information). Prior to the experimental session, participants' subjective levels of arousal were assessed with the SSS. Thereafter, we carried out a short training session. In a dimly lit, quiet room (similar across the four countries), participants rested their head on a chin support (the same for the four countries), 60 cm away from an LCD monitor of 1366 × 768 pixels of resolution. The experimenter explained the task together with two test trials, using sample pictures that were not included in the experimental set of images. Stimuli and timing parameters in the training session were equivalent to those in the actual experiment (see Fig. 1). Once the participant understood the task, we set up the eye tracker and ran a standard 9-point calibration. Participants held the mouse with both hands, and were instructed to press with their left or right index finger, on the left or right button of the mouse, respectively, depending on the answer. In those trials where participants did not answer, we assigned the maximum RT (following McKenna et al., 2006; Shahar et al., 2010). To avoid possible biases in response timing, the mapping between the right and the left buttons and the type of responses was balanced across participants and reversed for left-handed participants (n = 16). To answer the Likert scale in Stage 2, participants controlled the mouse with their dominant hand. After the first 70 trials, participants had a short break (less than 1 min). Participants had no control over the pace of the experiment, which lasted for approximately 30 min.

2.6. Statistical analysis

To test whether cross-cultural differences affected hazard perception, we first created a driving hazard ground truth based on the perceived breaking affordance from all drivers and scenes. We use the term *affordance* as defined by Norman (1988). That is, (braking) affordance refers to a specific relationship between road scene features (e.g., the presence of hazardous elements) and the driver behavior (considering [s]he is perusing a safe behavior). Thus, the braking affordance will not depend only on the actual characteristic of the road scene, but also on the driver's capabilities of acting, including his/her goals, beliefs, and past experiences. In the same way, we define "perceived braking affordance" as the subjective quantification of the affordance. The developing of the road hazard ground truth is an important step as there is an absence of a common set of reliable hazardous stimuli against which drivers' hazard perception can be compared and assessed (Martín de Diego et al., 2019). Hereinafter, we will use the terms "perceived hazard" or "perceived braking affordance" interchangeably. To create the road hazard ground truth, we used the assessment of the 218 international drivers and averaged the perceived braking affordance (scale from 0 to 6) from all drivers and driving scenes to obtain the mean rating of each picture. We then ordered the 140 pictures according to their level of breaking affordance from the lowest (0.54) to the highest (5.08) (average value = 2.48 ± standard deviation [SD] 1.17; median = 2.49) (see Fig. 2B). Fig. 2A presents the perceived braking affordance for the whole set of driving scenes for each country, as well as the average rating across the four countries. Once the scenes were graded, for descriptive purposes we calculated the mean and SD of the participants' RTs and gaze entropy depending on the country and the braking affordance.

To analyze variations in participants' RTs (natural log transformed) and gaze entropy over the 140 driving scenes, we estimated four multilevel models using maximum likelihood estimation. We used a step-up strategy to build the best model, from an "intercept-only model" (or empty model) to more complex models. The predictors included in the models (fixed effects) were the participants' *country* (participant-level predictor) and *braking affordance* (stimulus-level predictor). The effects of participants' *gender* and *driving experience* as well as the *visual saliency* of the images (understood as the conspicuity of objects in the visual environment, Itti and Koch, 2000) were also examined for their potential contribution. The only relevant predictor was *driving experience* in the RTs models. *Participant* and *picture* random effects were also

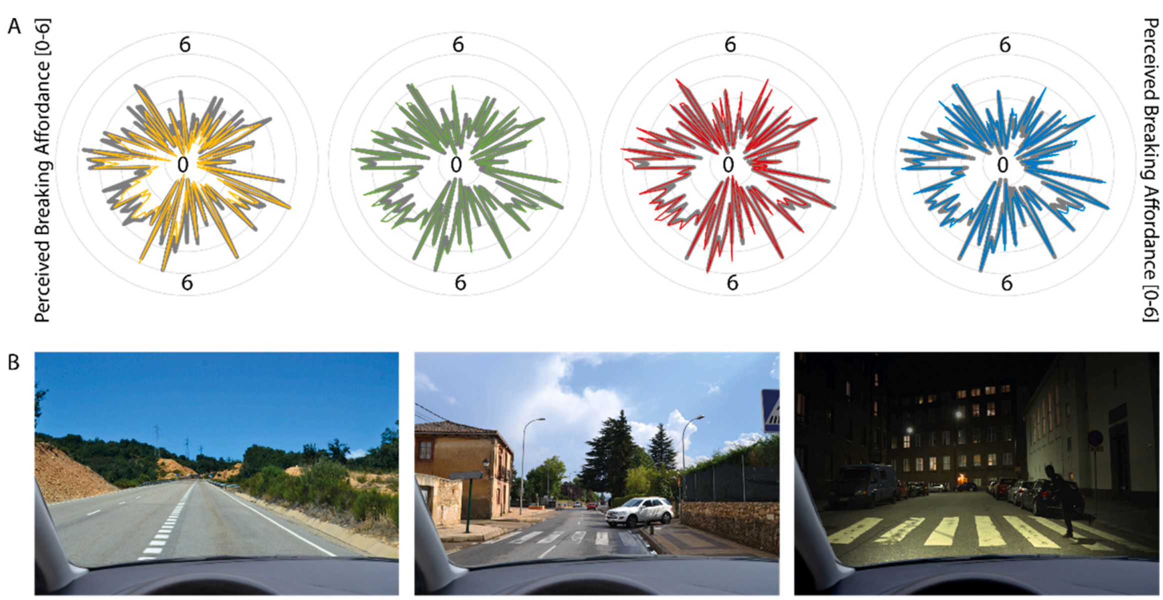

Fig. 2. A) Perceived braking affordance. Polar graphs present average perceived braking affordance from participants from each country (in yellow, Ukraine; in green, Italy; in red, Spain; in cyan, Sweden) for the whole set of driving scenes. The grey line represents the average rating of braking affordance given to each scene by drivers from the four countries (n = 218), our road hazard ground truth. In the vertical axis, the 7-point rating scale from the pole, 0 (not hazardous), to the outer circle, 6 (extremely hazardous). In the polar axis, clockwise, the rating for each driving scene, from picture 1 at 12:00 (average breaking affordance 1.18) to picture 140 at 11:59 (average breaking affordance 1.87). B) Examples of driving scenes. Pictures that were rating, from left to right, as having relative low ($\overline{X} = 0.54$), medium ($\overline{X} = 3.23$), and high ($\overline{X} = 5.08$) braking affordance for drivers from the four countries. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

included.

For both dependent variables (RTs and gaze entropy), we first fitted an "intercept-only model" (Model 1), with *participants* and *pictures* as random effects and without any predictor. Second, we fitted a "main effects model" (Model 2), including also participants' *country*, and *braking affordance* (*driving experience* in the case of RT too) into the model. Third, we fitted an "all interactions model" (Model 3) that also included all two-way interactions among *country* and *braking affordance* (and interactions among the *country* and *braking affordance* with *driving experience* in the case of RT). The final model was a restricted version that only included significant interactions (Model 4). For each model, we present parameter estimates and standard errors. Note that to assess the effect of *country*, we created three dummy variables, using being Swedish as the reference (coded as 0). To assess the effect of *driving experience*, we created a dummy variable with novice drivers (with 5 or less years of experience) as the reference (coded as 0).

To assess the improvement in model fit, we used a correction of the Akaike's Information Criterion for small samples (AICc, a lower AICc value indicates that the model fits the data better). Also, we dismissed models whose probability to minimize the information loss was below 0.05 (compared to the lowest AICc score) using the following formula:

$$P = e^{(AICc_{min} - AICc_i)/2}$$

Finally, we carried out a univariate 4 × 2 ANOVA in order to examine main and interaction effects of *country* and *driving experience* on driver braking behavior (number of brakes). Bonferroni correction for multiple comparisons was applied with a critical level set at $\alpha < 0.05$ and partial η^2 (η_p^2) was calculated to estimate the effect size.

3. Results

Descriptive statistics of RT, gaze entropy, and braking behavior depending on the participants' *country* and the *braking affordance* of the scenes are shown in Table 3.

3.1. Response times

In the "intercept-only model" (Model 1), the fixed intercept was 1.118 s, the expected RT (natural log transformed RT = 0.112) for an average participant on an average picture. Of the total variance (0.112), 19.6% was between *participants* and 20.4% was between *pictures*. That is, 95% of the participants would be expected to have an individual intercept between 0 and 0.41 (1.50 s) for an average item, and 95% of the pictures would be expected to have an intercept between 0 and 0.41 (1.51 s) for an average participant. In the "main effects model" (Model 2), the fixed effects of *braking affordance* in the driving scenes (-0.084) were significant. This parameter estimate represents the expected linear rate of decrease in RT for a one-unit increase in braking affordance. The fixed effects of *driving experience* (0.009) were not significant. The fixed effects of being *Ukrainian* (-0.105) were significant, but the fixed effects of being *Italian* (-0.044), and being *Spanish* (-0.047) were not. These fixed effects represent the expected difference in the intercept between Swedish drivers (0.364 or 1.44 s) and Ukrainian, Italian, and Spanish drivers, respectively. In the "all interactions model" (Model 3), the interaction terms *braking affordance* × *Italy*, *driving experience* × *Ukraine*, *driving experience* × *Italy*, and *driving experience* × *Spain* were not significant. They were therefore removed from the overall model. In the "significant interactions model" (Model 4), the fixed effects of *braking affordance* in the driving scenes (-0.087) and being *Ukrainian* (-0.133) showed significant estimates. The interactions between *driving experience* × *braking affordance* (0.014), *braking affordance* × *Ukraine* (0.011) and *braking affordance* × *Spain* (-0.021) were also significant. Models 1, 2, and 3 were dismissed as their probability to minimize information loss (when compared to Model 4) was below 0.05. Note that the effects of drivers' *gender* were examined in preliminary analysis, but they did not contribute significantly to the models (data not shown).

Table 4 details all parameter estimates, standard errors, and significant tests for fixed and random effects for the four models.

3.2. Gaze entropy (visual search strategy)

In the "intercept-only model" (Model 1), the fixed intercept was 3.312 bits, the expected gaze entropy for an average participant on an average picture. Of the total variance (0.326), 27% was between *participants* and 17% was between *pictures*. That is, 95% of the participants would be expected to have an individual intercept between 2.72 and 3.90 bits for an average item, and 95% of the pictures would be expected to have an intercept between 2.83 and 3.79 bits for an average participant. In the "main effects model" (Model 2), the fixed effects of the *braking affordance* (-0.035) and being *Spanish* (-0.249) were significant. The first parameter estimate represents the expected linear rate of decrease in gaze entropy for a one-unit increase in braking affordance. The second parameter estimate represents the expected difference in the intercept between Swedish (3.479 bits) and Spanish drivers. In the "all interactions model" (Model 3), the interaction terms *braking affordance* × *Italy* (0.032) and *braking affordance* × *Spain* (0.030) were significant. The interaction term *perceived braking affordance* × *Ukraine* (0.011) was not significant. It was therefore was removed from the overall model. AICc was bigger in the "significant interactions model" (Model 4). However, Model 4 is 0.559 times as probable as the Model 3 to be the best model. Models 1 and 2 were dismissed as their probability to minimize information loss (when compared to Model 3) were below 0.05. Note that the effects of drivers' *gender* and *driving experience*, as well as the *visual saliency* of the pictures were examined in preliminary analysis, but they did not contribute significantly to the models (data not shown).

Table 5 details all parameter estimates, standard errors, and significant tests for fixed and random effects for the four models.

Table 3
Means and standard deviations (in parenthesis) of response time (RT, sec) and gaze entropy (GE, bits) depending on the *country* and the *braking affordance* (BA, scale from 0 to 6). The braking affordance was calculated from all 218 drivers and driving scenes. For descriptive purposes, we clustered the scores of the 140 pictures from the lowest braking affordance (cluster 0-0.99) to the highest braking affordance (cluster 5–6). The BA column shows the ranges of scores for all clusters and the number of pictures that belong to each cluster. The percentage of brakes (%B) presents the braking behavior from all drivers in each country to the driving scenes in each cluster.

	Ukraine			Italy			Spain			Sweden		
	$n = 46$			$n = 53$			$n = 66$			$n = 53$		
BA	%B	RT	GE	%B	RT	GE	%B	RT	GE	%B	RT	GE
0–0.99 n = 15	6	1.11 (0.37)	3.37 (0.55)	6	1.14 (0.33)	3.44 (0.57)	9	1.22 (0.34)	3.15 (0.58)	6	1.21 (0.37)	3.46 (0.55)
1–1.99 n = 46	22	1.23 (0.38)	3.37 (0.52)	23	1.31 (0.37)	3.45 (0.54)	30	1.33 (0.36)	3.16 (0.54)	19	1.35 (0.38)	3.44 (0.48)
2–2.99 n = 25	61	1.19 (0.39)	3.44 (0.54)	64	1.28 (0.38)	3.49 (0.56)	72	1.25 (0.37)	3.24 (0.56)	51	1.38 (0.40)	3.49 (0.49)
3–3.99 n = 39	89	1.03 (0.36)	3.26 (0.60)	92	1.06 (0.34)	3.38 (0.58)	93	1.02 (0.31)	3.10 (0.60)	86	1.14 (0.39)	3.31 (0.59)
4–4.99 n = 14	94	0.91 (0.32)	3.22 (0.60)	96	0.96 (0.32)	3.38 (0.58)	96	0.94 (0.28)	3.08 (0.58)	89	0.97 (0.33)	3.26 (0.55)
5–6 n = 1	96	0.89 (0.30)	3.06 (0.56)	96	1.06 (0.33)	3.20 (0.60)	100	1.06 (0.29)	2.62 (0.63)	91	1.04 (0.29)	3.01 (0.58)

Table 4
Multilevel model parameter estimates and standard errors (in parenthesis) for response times (RTs, natural log transformed) in seconds depending on the participants' *driving experience* and *country* and the *braking affordance* in the driving scenes ($n = 218$). The fixed effects of *braking affordance* represent the expected linear rate of decrease in RT for a one-unit increase in braking affordance. The fixed effects of *driving experience* represent the differences in the intercept between novice (the reference) and experienced drivers in RT. The fixed effects of being *Ukrainian*, *Spanish*, and *Italian* represent the differences in the intercept between Ukrainians, Spaniards, and Italians compared to Swedes (the reference). The fixed effects of the interactions between *driving experience* and *braking affordance* represent the differences in the slopes between experienced and novice drivers. The fixed effects of the interactions between *braking affordance* and being *Ukrainian*, *Spanish*, and *Italian* represent the differences in the slopes between Swedes and Ukrainian, Spanish, and Italian drivers, respectively.

	Model 1	Model 2	Model 3	Model 4
Fixed effects				
Intercept	0.112 (0.016)**	0.364 (0.039)**	0.351 (0.052)**	0.371 (0.035)**
Driving experience	–	0.009 (0.021)	−0.008 (0.051)	−0.025 (0.022)
Braking affordance	–	−0.084 (0.008)**	−0.091 (0.008)**	−0.087 (0.008)**
Ukraine	–	−0.105 (0.031)*	−0.107 (0.053)	−0.133 (0.032)**
Italy	–	−0.044 (0.030)	−0.023 (0.053)	−0.044 (0.030)
Spain	–	−0.047 (0.028)	0.024 (0.052)	0.004 (0.028)
Driving experience × Braking affordance	–	–	0.016 (0.003)**	0.014 (0.002)**
Driving experience × Ukraine	–	–	−0.056 (0.069)	–
Driving experience × Italy	–	–	−0.060 (0.066)	–
Driving experience × Spain	–	–	−0.031 (0.062)	–
Braking affordance × Ukraine	–	–	0.015 (0.004)**	0.011 (0.003)*
Braking affordance × Italy	–	–	0.006 (0.003)	–
Braking affordance × Spain	–	–	−0.017 (0.003)**	−0.021 (0.002)**
Variance components				
Random intercept variance	0.022 (0.002)**	0.020 (0.002)**	0.020 (0.002)**	0.020 (0.002)**
Random picture variance	0.023 (0.003)**	0.013 (0.001)**	0.013 (0.001)**	0.013 (0.001)**
Residual variance	0.067 (0.0005)**	0.067 (0.0005)**	0.067 (0.0005)**	0.067 (0.0005)**
AICc	5,583.71	5,503.69	5,401.88	5,396.60

Note. Model 1 = Intercept-only model; Model 2 = Main effects model; Model 3 = All interactions model; Model 4 = Significant interactions model; AICc = corrected Akaike's Information Criterion.
* $p < 0.05$.
** $p < 0.001$.

Figure 3 displays the expected fixed effects of *braking affordance* in the driving scenes for the four *countries* on RTs (Model 4) and gaze entropy (Model 3). Response times decreased with increased *braking affordance*. The rate of decrease in RTs over increasing breaking affordance differed between experienced and novice drivers, and among *countries*. Novice drivers showed a faster rate of decrease in RT. Regarding differences among countries, Ukrainians showed a slower rate of decrease in RT compared to Swedes, and Spaniards, a faster rate of decrease. Gaze entropy also decreased with increased *braking affordance*, and the rate of change was different among *countries*. Spaniards and Italians showed a slower rate of decrease.

Finally, mean braking behavior for the 140 driving scenes pictures differed among the four *countries*. The main effect of *country* on the number of brakes was significant, $F(3, 214) = 11.50$, $p < 0.001$, $\eta_p^2 = 0.14$. Post hoc comparisons showed that drivers from Italy ($\bar{X} = 77.68$, $SD = 11.44$) braked more frequently than drivers from

Table 5
Multilevel model parameter estimates and standard errors (in parenthesis) for gaze entropy (visual search strategy) depending on the participants' *country* and the *braking affordance* in the driving scenes ($n = 218$). The fixed effects of *braking affordance* represent the expected linear rate of decrease in gaze entropy for a one-unit increase in braking affordance. The fixed effects of being *Ukrainian*, *Spanish*, and *Italian* represent the differences in the intercept between Ukrainians, Spaniards, and Italians compared to Swedes (the reference). The fixed effects of the interactions between *braking affordance* and being *Ukrainian*, *Spanish*, and *Italian* represent the differences in the slopes between Swedes and Ukrainian, Spanish, and Italian drivers, respectively.

	Model 1	Model 2	Model 3	Model 4
Fixed effects				
Intercept	3.312 (0.028)**	3.479 (0.060)**	3.528 (0.060)**	3.515 (0.060)**
Braking affordance	–	−0.035 (0.017)*	−0.054 (0.017)*	−0.049 (0.017)*
Ukraine	–	−0.060 (0.055)	−0.089 (0.057)	−0.061 (0.055)
Italy	–	0.031 (0.053)	−0.049 (0.055)	−0.036 (0.055)
Spain	–	−0.249 (0.051)**	−0.326 (0.053)**	−0.313 (0.052)**
Braking affordance × Ukraine	–	–	0.011 (0.006)	–
Braking affordance × Italy	–	–	0.032 (0.005)**	0.027 (0.005)**
Braking affordance × Spain	–	–	0.030 (0.005)**	0.026 (0.005)**
Variance components				
Random intercept variance	0.087 (0.008)**	0.074 (0.007)**	0.074 (0.007)**	0.074 (0.007)**
Random picture variance	0.057 (0.006)**	0.055 (0.006)**	0.055 (0.006)**	0.055 (0.006)**
Residual variance	0.180 (0.001)**	0.180 (0.001)**	0.180 (0.001)**	0.180 (0.001)**
AICc	35,917.62	35,886.57	35,849.71	35,850.87

Note. Model 1 = Intercept-only model; Model 2 = Main effects model; Model 3 = All interactions model; Model 4 = Significant interactions model; AICc = corrected Akaike's Information Criterion.
* $p < 0.05$.
** $p < 0.001$.

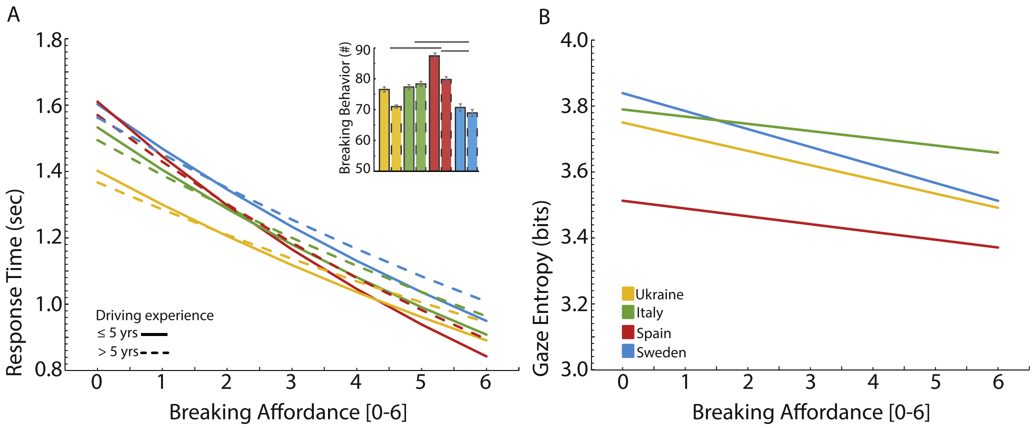

Fig. 3. (A) Expected effects of the braking affordance present in the driving scenes for Ukrainian, Italian, Spanish, and Swedish drivers on response times, the promptness with which drivers decided to brake or to keep driving (in seconds). Solid lines represent novice drivers. Dash lines represent experienced drivers. Inset. Mean driver braking behavior for the 140 driving scenes for novice and experienced drivers from the four countries (Ukraine, $n = 46$; Italy, $n = 53$; Spain, $n = 66$; Sweden, $n = 53$). Error bars represent standard errors of the mean. Significantly differing countries (main effects) have a line above, connecting them both (corrected p-values < 0.05). (B) Expected effects of the braking affordance present in the driving scenes for Ukrainian, Italian, Spanish, and Swedish drivers on gaze entropy, drivers' visual search strategies (in bits). Colors as in Fig. 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Sweden ($\overline{X} = 69.26$, $SD = 16.84$), corrected $p < 0.05$. Drivers from Spain ($\overline{X} = 83.83$, $SD = 14.14$) braked more frequently than drivers from Ukraine ($\overline{X} = 74.97$, $SD = 11.09$) and Sweden, corrected *ps.* < 0.05. The main effect of *driving experience* on the number of brakes was not significant, $p > 0.05$. Novice ($\overline{X} = 79.63$, $SD = 13.90$) and experienced drivers ($\overline{X} = 73.96$, $SD = 14.93$) did not differ in their braking behavior. The interaction term *country* × *driving experience* was not significant as well, $p > 0.05$ (see inset in Fig. 3A).

4. Discussion

In cross-cultural studies, the different methodologies used to measure road hazard perception have made it difficult to assess the actual influence of culture and the characteristics of this influence. Here, we studied national differences on road hazard perception among young drivers from four European countries (Ukraine, Italy, Spain, and Sweden). For the first time, an identical set up for multimodal assessment, which included behavioral and physiological indices, was used. The results show a general decrease in RT over increasing levels of braking affordance. Drivers from Ukraine, the participating country with the highest death rates, were the fastest to respond, whereas drivers from Sweden, the country with the lowest death rates, were the slowest. Moreover, as the braking affordance increased, the gaze entropy decreased, meaning that the visual search strategy became less erratic. The gaze entropy data furthermore showed that Spanish drivers had the most structured visual search strategy over all levels of braking affordance, whereas the most rigid visual search patterns were found among Italian drivers.

Even though the countries included in the study meet the best practice criteria on several key risk factors for road safety (as defined by World Health Organization, 2018), such as drinking and driving or seat-belt laws, they differ in other aspects, such as the rate of traffic-related deaths or the amount of yearly hours spent in road congestion (see Table 1 for more details about their differences). The present study suggests that hazard perception can be defined cross-culturally, with cultural factors modulating RTs, braking responses, and visual search strategies.

4.1. Response times and breaking responses

To implement a response to reduce the likelihood of having an accident, a driver first has to detect the potential hazard and, then, to recognize the potential hazard as a true hazard (Huang and Winston, 2011). This chain of decision-making events usually takes about two seconds (Olson and Sivak, 1986; Summala, 2000). Our results suggest that this process can be even faster. With a limit for collection of RT at two seconds, the results still show a linear decrease in RTs over increasing levels of breaking affordance, with expected RTs below one second at the greatest level of breaking affordance. This result is independent of the drivers' country of origin and confirms previous findings on hazard perception using similar experimental paradigms (Ho et al., 2001; Huestegge et al., 2010; Serrano et al., 2014; Torres et al., 2017). Once a driver has recognized a potential hazardous target, there is no need to delay the response. However, in low risk situations (where hazards are less clear), the uncertainty of action outcomes seems to promote an exploratory behavior (Doya, 2008). For example, participants would keep searching for potential hazards, which would delay driver response. This explanation is also supported by the modulation of the visual search strategy by breaking affordance (see section 4.2).

Regarding cross-cultural differences, Ukrainian and Swedish drivers show opposite trends in their RTs. This is in line with our hypothesis. However, contrary to our assumption, Ukrainian drivers are the fastest to respond, while the Swedish drivers are the slowest. Lim and colleagues (2013) suggested that actual experience with road hazards may lead to two different outcomes: It could either sensitize drivers, who would be quicker and more accurate in detecting hazards, or it would desensitize drivers, who would be slower and less accurate in detecting hazards. In light of our findings, it seems that there is a third possible outcome. Ukrainian drivers are the fastest in making their decisions, but the decisions would not be necessarily the safest, looking at the rate of braking. This result is coherent with the fact that there is a higher probability of being involved in a traffic accident on Ukrainian roads (World Health Organization, 2018). That is, Ukrainian drivers could have a more urgent than evaluative response pattern (as defined by Megías et al., 2011). Moreover, by spending less time assessing the traffic situation, drivers could increase the time devoted to performing other tasks (Ho et al., 2001). This behavior can denote an adaptive response to the overall Ukrainian hostile traffic climate (Zahreba, 2017). Not only did the Swedish drivers have the slowest responses, but also the lowest frequency of braking behavior. A potential reason to explain these results is the optimism bias, a systematic misperception about having lower probabilities to experience negative events (DeJoy, 1989). This phenomenon has already been reported among Nordic drivers (Nordfjærn et al., 2011). Briefly, since most Swedish drivers do not experience many adverse consequences while driving, they judge negative consequences as unlikely (Andersson, 2007). This bias can interfere with road safety by affecting many tactical and strategic safety decisions. It is well known that drivers who feel themselves to be relatively immune may decide not to brake (Svenson et al., 1985). The braking rate was similar between Swedish and Ukrainian drivers, which offers additional support to the hypothesis that both optimistic and angry drivers might interfere with road safety (Islam and Mannering, 2020; Stephens and Ohtsuka, 2014). However, there is a large difference in prevalence of traffic accident death rates between the countries, where the Swedish roads are among the safest in the world, whereas Ukraine has the least safe roads in Europe. In line with our hypothesis, Spanish and Italian drivers showed similar RTs and tendencies across the levels of braking affordance, positioned in between the Ukrainian and Swedish drivers.

Although it was not a primary aim of the current study, it is worth noting that the linear decrease in RTs over increasing levels of braking affordance was influenced by driving experience. Novice drivers would be faster in making their decisions while assessing driving scenes with higher levels of braking affordance. This finding contradicts some previous results where experienced drivers were faster (Huestegge and Böckler, 2016). Differences in the set of stimuli and/or the definition of driving experience could explain such inconsistencies. Additionally, a driver's gender did not influence RTs, which is in accordance with previous research using similar research paradigms (Tūskė et al., 2019). However, the generalizability of these specific findings are somewhat limited by the unequal distributions in driving experience and gender among the samples from the four countries. Further research is needed to investigate whether there are specific gender and/or driving experience-based differences across different countries.

4.2. Visual search strategy

Our results show that drivers' visual search strategy are modulated by the braking affordance of the driving scenes. The larger the perceived hazard, the lower the gaze dispersion (entropy). These findings confirm the original phenomenon reported by Underwood and colleagues (Chapman and Underwood, 1998; Crundall et al., 1999; Underwood et al., 2011, 2005). However, studying the phenomenon cross-culturally, we are the first to show modulation of cultural factors on gaze dispersion. That is, encountering a hazardous event on the road elicits a general reduction of drivers' gaze dispersion (Savage et al., 2020). As explained by Underwood and colleagues, "when a hazardous area of the scene has been potentially identified, it is important that information in this region is processed in depth. There may be advantages to monitoring that location for further unfolding of events" (Underwood et al., 2011). On the contrary, when a clear risky element is not present, as happens in the low braking affordance pictures, it is plausible to assume that drivers adopt a strategy able to maximize the chances of spotting a potential hazard (i.e., an increase in gaze dispersion).

Regarding cross-cultural differences, contrary to our hypothesis, Ukrainian and Swedish drivers do not differ in their levels and tendencies of gaze dispersion across the driving scenes. Instead, Italian and Spanish drivers show the most interesting results. Italian drivers have a more rigid visual strategy: their gaze dispersion was the highest and it was almost constant across the different levels of breaking affordance. This search pattern, together with a higher breaking rate, is consistent with a defensive driving style (e.g., continually controlling surrounding traffic). It is plausible that Italian drivers display a unique visual search strategy as a behavioral adaptation caused by the national traffic climate. Italian are less rule-oriented and obedient than, for example, Spanish drivers (de Oña et al., 2014), and likely, Swedish drivers. Furthermore, considering the overall Italian traffic density (one of the highest in EU; European Automobile Manufacturers' Association, 2018; International Traffic Safety Data and Analysis Group, 2019) and the frequent violations of the traffic laws by cyclists and pedestrians (Fraboni et al., 2018), a constant and wide visual search pattern can represent an optimal adaptation to prevent traffic accidents. Spanish drivers showed the lowest level of gaze dispersion, compared to drivers from all other countries, indicating a more structured visual search strategy. A straightforward reason for these results is not apparent. However, a plausible explanation might be related to the experimental paradigm we used. Since 2013, the Spanish National Department of Traffic is testing and promoting the introduction of hazard perception tests for trainee drivers (Castro et al., 2014). If we consider the age of our Spanish sample (< 24 years old), it is possible that they were familiar with the procedure. Previous exposure to a similar test might have biased their visual search strategy (Kahana-Levy et al., 2019).

The effects of braking affordance on gaze entropy would not be influenced by driving experience, nor by drivers' gender. Whatever the reason for the lack of effects of these factors in our study, it is worth noting that they replicate those observed in previous studies regarding gaze metrics while examining risky driving scenes (Lim et al., 2013) and a driver's visual scanning under complex road situations (Bao and Boyle, 2009). Again, examining the influence of such drivers' characteristics was not a primary aim within the current study, and therefore future studies should address this issue.

4.3. Neural underpinnings of the effect of cultural differences in road hazard perception

In the last decade, studies have shown that cultural experiences, through either practice or observation, might affect brain functioning (Crafa and Nagel, 2020; Park and Huang, 2010). Then, it is quite reasonable to hypothesize that sustained exposure to particular traffic experiences (e.g., road traffic violations) and behavioral practices (e.g., aggressive behaviors while driving) might affect also brain functioning. Indeed, studies using neuroimaging have shown that participants from different cultures respond to similar situations in a different manner, showing specific activity patterns (Shkurko, 2020). This result would also happen while completing risk taking tasks (Qu et al., 2019). In the same line, culture-based differences have been observed in the ventral visual cortex [VVC] when participants passively observed still pictures (for a review on this topic, see Park and Huang, 2010). The VVC is an area highly associated with the visual and perceptual processing, including visual search strategies (Jo et al., 2019). Thus, it is plausible to think that functional differences in this area may explain, in part, the differences observed here in visual search strategies and RT, due to drivers' culture. Future cross-cultural works in road hazard perception may shed some further light on this issue, integrating neuroimaging and eye tracking methodologies (Kim et al., 2020).

4.4. Limitations of the study

The current study is the first large-scale attempt to compare hazard perception among four European countries, using the identical multimodal assessment methodology and technology. However, the overall results should be viewed in the context of three shortcomings related to the experimental procedure. First, although the sample size per country was similar to previous studies that have used eye movements (Lim et al., 2013), it would have been preferable to have a larger sample size in order to reach a nation-wide conclusion and further studies are warranted. Second, we assigned the maximum RT to those trials where the driver failed to respond (4.2% of the trials). Thus, the RT was a measure of how quickly drivers responded, relying on the assumption that everyone would have responded eventually, even if the picture had disappeared. This strategy has been used before in similar studies (for a discussion on this topic, see Shahar et al., 2010). Third, we used still images to study road hazard perception, and not video-clips, which are considered a more ecological instrument. It is, however, worth noting that tests containing static images present some advantages compared to those that use dynamic images: (1) the unambiguousness of the onset of the hazard stimulus, (2) the lower cost of the development and administration of the test, and (3) the efficiency in time and number of stimuli presented (Moran et al., 2019; Scialfa et al., 2012; Tūskė et al., 2019).

4.5. Future directions

Outside of the realm of road safety, when studying human behavior, normed and cross-cultural validated stimuli are quite common (e.g., International Affective Picture System [Lang et al., 1997]). Unfortunately, in hazard perception studies, a set of universal validated traffic stimuli has not been developed yet (Martín de Diego et al., 2019). Actually, hazard perception stimuli are often created idiosyncratically for each study in order to minimize the potential influence of sceneries that are unfamiliar, either culturally or environmentally, to the tested population. In addition, hazard perception tests developed in a particular country may not be similarly effective when drivers from other cultures undertake them (e.g., Bazilinskyy et al., 2020). Moreover, the question of whether a “western bias” (Henrich et al., 2010) might be present also in road hazard perception studies is currently open. Therefore, the need of creating a cross-culturally validated set of stimuli should be fulfilled and our study represents a first approximation. Future works should pursue the development and validation of a set of road traffic stimuli that can be used also in under-represented cultures. Additionally, considering the simplicity and economy of the implemented method, this set of stimuli could be used for driver evaluation and training in low-income countries, where the lack of infrastructure to support the use of advanced technologies in traffic research is one of the main limitations in testing national road safety policies (Perego et al., 2018).

4.6. Conclusions

Road traffic accidents represent a major health issue worldwide. We have shown that regional cultures may influence drivers’ interpretation of a traffic situation and, consequently, the correct identification of potentially hazardous situations. Here, we compared road hazard perception among Ukrainian, Italian, Spanish, and Swedish young drivers, using the same multimodal assessment methodology. Overall, our results differentiated between drivers’ origin with regard to all included measures, behavioral responses as well as visual search strategies. This could be explained by differences in traffic cultures and legislations, or by differences in other sociocultural factors. Moreover, we reported that the levels of braking affordance influence RTs as well as visual search strategies, coherently across the four countries. That is, independently of drivers’ country of origin, drivers react quickly and have a more structured visual exploration when presented to a high risk situation, whereas they reacted slower and showed a more dispersed visual exploration when presented to a low risk situation. Finally, our results also support the idea that a multimodal research approach is possible for mass driver testing. As gaze-based metrics provide a valuable source of information, their inclusion might in part overcome the criticism of the simplicity of RT tests used as a means of evaluating hazard perception skills (for a recent discussion on this topic, see Horswill et al., 2020).

Funding

This work was funded by the Ramón y Cajal fellowship program from the Spanish State Research Agency (RYC-2015-17483) awarded to LLDS. Additional support was obtained from the Unit of Excellence on Brain, Behavior, and Health (SC2) funded by the Excellence actions program, from the research group Information Processing Circuits and Systems (TIC-117) both from the University of Granada, and from the Faculty of Human Sciences (Mid Sweden University). The Erasmus + for teaching mobility and the Erasmus + action KA107 provided mobility support to LLDS, CDP, and JMM. The work was carried out within the scope of the research program “Dipartimenti di Eccellenza”, which was supported by a grant from MIUR to the Department of General Psychology, University of Padua. These organizations had no role in the design or conduct of this research.

CRediT authorship contribution statement

Leandro L. Di Stasi: Conceptualization, Investigation, Visualization, Writing - original draft, Writing - review & editing, Funding acquisition, Supervision. **Carolina Diaz-Piedra:** Methodology, Formal analysis, Writing - review & editing. **José M. Morales:** Software, Data curation, Investigation. **Anton Kurapov:** Investigation. **Mariaelena Tagliabue:** Investigation, Writing - review & editing. **Anna Bjärtå:** Investigation, Writing - review & editing. **Alberto Megias:** Resources, Writing - review & editing. **Jens Bernhardsson:** Resources, Writing - review & editing. **Svitlana Paschenko:** Resources. **Samuel Romero:** Resources. **Antonio Cándido:** Writing - review & editing. **Andrés Catena:** Resources, Writing - review & editing.

Declaration of Competing Interest

None.

Acknowledgements

The authors would like to thank Prof. G. Vidotto, Department of General Psychology, University of Padova, for his support during data collection, and O. Shevchuk, Taras Shevchenko National University of Kyiv, for her help during data collection. In addition, the authors acknowledge C. Posey, Mind, Brain, and Behavior Research Center-CIMCYC, University of Granada, for her comments and suggestions on the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.aap.2020.105785.

References

Ābele, L., Haustein, S., Møller, M., Martinussen, L.M., 2018. Consistency between subjectively and objectively measured hazard perception skills among young male drivers. Accid. Anal. Prev. 118, 214–220. https://doi.org/10.1016/j.aap.2018.02.022.

Adminaité-Fodor, D., Heilpern, C., Jost, H., 2019. Ranking EU Progress on Road Safety 13th Road Safety Performance Index Report. European Transport Safety Council, Brussels, Belgium.

Andersson, H., 2007. Willingness to pay for road safety and estimates of the risk of death: evidence from a Swedish contingent valuation study. Accid. Anal. Prev. 39, 853–865.

Antov, D., Banet, A., Barbier, C., Bellet, T., Bimpeh, Y., Boulanger, A., et al., 2012. European Road Users’ Risk Perception and Mobility: the SARTRE 4 Survey. Institut français des sciences et technologies des transports, de l’aménagement et des réseaux (IFSTTAR).

Baltruschat, S., Cándido, A., Megías, A., Maldonado, A., Catena, A., 2020. Risk proneness modulates the impact of impulsivity on brain functional connectivity. Hum. Brain Mapp. 41, 943–951. https://doi.org/10.1002/hbm.24851.

Bao, S., Boyle, L.N., 2009. Age-related differences in visual scanning at median-divided highway intersections in rural areas. Accid. Anal. Prev. 41, 146–152. https://doi.org/10.1016/j.aap.2008.10.007.

Bazilinskyy, P., Eisma, Y.B., Dodou, D., Winter de, J.C.F., 2020. Risk perception: a study using dashcam videos and participants from different world regions. Traffic Inj. Prev. 21, 347–353. https://doi.org/10.1080/15389588.2020.1762871.

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436.

Carson, J., Adminaité-Fodor, D., Jost, G., 2020. Ranking EU Progress on Road Safety. 14th Road Safety Performance Index Report. European Transport Safety Council, Brussels, Belgium.
Cassarino, M., Murphy, G., 2018. Reducing young drivers' crash risk: are we there yet? An ecological systems-based review of the last decade of research. Transp. Res. Part F Traffic Psychol. Behav. 56, 54–73. https://doi.org/10.1016/j.trf.2018.04.003.
Castro, C., Padilla, J.L., Roca, J., Benítez, I., García-Fernández, P., Estévez, B., López-Ramón, M.F., Crundall, D., 2014. Development and validation of the spanish hazard perception test. Traffic Inj. Prev. 15, 817–826. https://doi.org/10.1080/15389588.2013.879125.
Chapman, P.R., Underwood, G., 1998. Visual search of driving situations: danger and experience. Perception 27, 951–964. https://doi.org/10.1068/p270951.
Connor, J., Norton, R., Ameratunga, S., Robinson, E., Civil, I., Dunn, R., Bailey, J., Jackson, R., 2002. Driver sleepiness and risk of serious injury to car occupants: population based case control study. BMJ 324, 1125. https://doi.org/10.1136/bmj.324.7346.1125.
Crafa, D., Nagel, S.K., 2020. Traces of culture: the feedback loop between behavior, brain, and disorder. Transcult. Psychiatry 57, 387–407.
Crundall, D., Underwood, G., 2011. Visual attention while driving. Handbook of Traffic Psychology. Elsevier Academic Press, San Diego, CA, US, pp. 137–148. https://doi.org/10.1016/B978-0-12-381984-0.10011-6.
Crundall, D., Underwood, G., Chapman, P., 1999. Driving experience and the functional field of view. Perception 28, 1075–1087.
de Oña, J., de Oña, R., Eboli, L., Forciniti, C., Mazzulla, G., 2014. How to identify the key factors that affect driver perception of accident risk. A comparison between Italian and Spanish driver behavior. Accid. Anal. Prev. 73, 225–235. https://doi.org/10.1016/j.aap.2014.09.020.
de Winter, J.C.F., Dodou, D., 2016. National correlates of self-reported traffic violations across 41 countries. Pers. Individ. Dif. 98, 145–152. https://doi.org/10.1016/j.paid.2016.03.091.
DeJoy, D.M., 1989. The optimism bias and traffic accident risk perception. Accid. Anal. Prev. 21, 333–340.
Diaz-Piedra, C., Gomez-Milan, E., Di Stasi, L.L., 2019. Nasal skin temperature reveals changes in arousal levels due to time on task: an experimental thermal infrared imaging study. Appl. Ergon. 81, 102870. https://doi.org/10.1016/j.apergo.2019.06.001.
Díaz-Román, A., Megías, A., Diaz-Piedra, C., Catena, A., Di Stasi, L.L., 2015. Seeing red on the road. Perception 44, 87–92. https://doi.org/10.1068/p7811.
Directorate-General for Trade European Commission, 2020. Trade in Goods With Ukraine (No. Units A4/G2).
Doya, K., 2008. Modulators of decision making. Nat. Neurosci. 11, 410–416.
Ehsani, J.P., Seymour, K.E., Chirles, T., Kinnear, N., 2020. Developing and testing a hazard prediction task for novice drivers: a novel application of naturalistic driving videos. J. Safety Res. https://doi.org/10.1016/j.jsr.2020.03.010.
European Automobile Manufacturers' Association, (ACEA), 2018. ACEA Report: Vehicles in Use–Europe 2018. Brussels, Belgium.
Eurostat, 2018. First Residence Permits Issued in the EU Member States Remain Above 3 Million in 2017 (No. 166/2018). Brussels, Belgium.
Fraboni, F., Marín Puchades, V., De Angelis, M., Pietrantoni, L., Prati, G., 2018. Red-light running behavior of cyclists in Italy: an observational study. Accid. Anal. Prev. 120, 219–232. https://doi.org/10.1016/j.aap.2018.08.013.
Grayson, G.B., Sexton, B.F., 2002. The Development of Hazard Perception Testing (No. TRL558). TRL Limited, Crowthorne, UK.
Harris, R.L., Glover, B.J., Spady, A.A., 1986. Analytical Techniques of Pilot Scanning Behavior and Their Application (Technical Paper No. NASA-Technical Paper-2525). NASA Langley Research Center, Hampton, VA, USA.
Henrich, J., Heine, S.J., Norenzayan, A., 2010. The weirdest people in the world? Behav. Brain Sci. 33, 61–83. https://doi.org/10.1017/S0140525X0999152X.
Hill, T., Sullman, M.J.M., Stephens, A.N., 2019. Mobile phone involvement, beliefs, and texting while driving in Ukraine. Accid. Anal. Prev. 125, 124–131. https://doi.org/10.1016/j.aap.2019.01.035.
Ho, G., Scialfa, C.T., Caird, J.K., Graw, T., 2001. Visual search for traffic signs: the effects of clutter, luminance, and aging. Hum. Factors 43, 194–207. https://doi.org/10.1518/001872001775900922.
Hoddes, E., Dement, W., Zarcone, V., 1972. The development and use of Stanford Sleepiness Scale (SSS). Psychophysiology 10, 431–436.
Horswill, M.S., McKenna, F., 2004. Drivers' hazard perception ability: situation awareness on the road. In: Banbury, S., Tremblay, S. (Eds.), A Cognitive Approach to Situation Awareness: Theory and Application. Ashgate Publishing, Ltd., UK, pp. 55–175.
Horswill, M.S., Anstey, K.J., Hatherly, C.G., Wood, J.M., 2010. The crash involvement of older drivers is associated with their hazard perception latencies. J. Int. Neuropsychol. Soc. 16, 939–944. https://doi.org/10.1017/S135561771000055X.
Horswill, M.S., Hill, A., Jackson, T., 2020. Scores on a new hazard prediction test are associated with both driver experience and crash involvement. Transp. Res. Part F Traffic Psychol. Behav. 71, 98–109. https://doi.org/10.1016/j.trf.2020.03.016.
Huang, P., Winston, F.K., 2011. Young drivers. Handbook of Traffic Psychology. Elsevier, pp. 315–338.
Huestegge, L., Böckler, A., 2016. Out of the corner of the driver's eye: peripheral processing of hazards in static traffic scenes. J. Vis. 16, 11. https://doi.org/10.1167/16.2.11.
Huestegge, L., Skottke, E.-M., Anders, S., Müsseler, J., Debus, G., 2010. The development of hazard perception: dissociation of visual orientation and hazard processing. Transp. Res. Part F Traffic Psychol. Behav. 13, 1–8. https://doi.org/10.1016/j.trf.2009.09.005.
International Traffic Safety Data and Analysis Group, 2019. Road Safety Annual Report 2019 - Italy (ANNUAL REPORT). Organisation for Economic Co-operation and Development (OECD), Paris, France.
Islam, M., Mannering, F., 2020. A temporal analysis of driver-injury severities in crashes involving aggressive and non-aggressive driving. Anal. Methods Accid. Res. 27, 100128. https://doi.org/10.1016/j.amar.2020.100128.
Itti, L., Koch, C., 2000. A saliency-based search mechanism for overt and covert shifts of visual attention. Vision Res. 40, 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7.
James, S.L., Lucchesi, L.R., Bisignano, C., Castle, C.D., Dingels, Z.V., Fox, J.T., et al., 2020. Morbidity and mortality from road injuries: results from the Global Burden of Disease Study 2017. Inj. Prev. https://doi.org/10.1136/injuryprev-2019-043302.
Jo, H.-G., Kellermann, T., Baumann, C., Ito, J., Schulte Holthausen, B., Schneider, F., Grün, S., Habel, U., 2019. Distinct modes of top-down cognitive processing in the ventral visual cortex. NeuroImage 193, 201–213. https://doi.org/10.1016/j.neuroimage.2019.02.068.
Kahana-Levy, N., Shavitzky-Golkin, S., Borowsky, A., Vakil, E., 2019. The effects of repetitive presentation of specific hazards on eye movements in hazard perception training, of experienced and young-inexperienced drivers. Accid. Anal. Prev. 122, 255–267. https://doi.org/10.1016/j.aap.2018.09.033.
Kim, H.-C., Jin, S., Jo, S., Lee, J.-H., 2020. A naturalistic viewing paradigm using 360° panoramic video clips and real-time field-of-view changes with eye-gaze tracking. NeuroImage 216, 116617. https://doi.org/10.1016/j.neuroimage.2020.116617.
Lajunen, T., Summala, H., 2003. Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses. Transp. Res. Part F Traffic Psychol. Behav. 6, 97–107. https://doi.org/10.1016/S1369-8478(03)00008-1.
Lang, P.J., Bradley, M.M., Cuthbert, B.N., 1997. International Affective Picture System (IAPS): Technical Manual and Affective Ratings. NIMH Center for the Study of Emotion and Attention, Gainesville, FL, US.
Lee, Y.M., Miller, K., Crundall, D., Sheppard, E., 2020. Cross-cultural effects on detecting multiple sources of driving hazard: evidence from the deceleration detection flicker test. Transp. Res. Part F Traffic Psychol. Behav. 69, 222–234. https://doi.org/10.1016/j.trf.2020.01.013.
Lim, P.C., Sheppard, E., Crundall, D., 2013. Cross-cultural effects on drivers' hazard perception. Transp. Res. Part F Traffic Psychol. Behav. 21, 194–206. https://doi.org/10.1016/j.trf.2013.09.016.
Maldonado, A., Torres, M.A., Catena, A., Cándido, A., Megías-Robles, A., 2020. From riskier to safer driving decisions: the role of feedback and the experiential automatic processing system. Transp. Res. Part F Traffic Psychol. Behav. 73, 307–317. https://doi.org/10.1016/j.trf.2020.06.020.
Martín de Diego, I., Siordia, O.S., Fernández-Isabel, A., Conde, C., Cabello, E., 2019. Subjective data arrangement using clustering techniques for training expert systems. Expert Syst. Appl. 115, 1–15. https://doi.org/10.1016/j.eswa.2018.07.058.
McKenna, F.P., Horswill, M.S., Alexander, J.L., 2006. Does anticipation training affect drivers' risk taking? J. Exp. Psychol. Appl. 12, 1–10. https://doi.org/10.1037/1076-898X.12.1.1.
Meesmann, U., Torfs, K., Nguyen, H., Van den Berghe, W., 2018. Do We Care About Road Safety? Key Findings From the ESRA1 Project in 38 Countries. ESRA Project (E-Survey of Road Users' Attitudes). (No. 2018- R- 02- EN). European Survey of Road users' safety Attitudes project, Brussels, Belgium.
Megías, A., Maldonado, A., Cándido, A., Catena, A., 2011. Emotional modulation of urgent and evaluative behaviors in risky driving scenarios. Accid. Anal. Prev. 43, 813–817. https://doi.org/10.1016/j.aap.2010.10.029.
Megías, A., Navas, J.F., Petrova, D., Cándido, A., Maldonado, A., Garcia-Retamero, R., Catena, A., 2015. Neural mechanisms underlying urgent and evaluative behaviors: an fMRI study on the interaction of automatic and controlled processes. Hum. Brain Mapp. 36, 2853–2864. https://doi.org/10.1002/hbm.22812.
Megías, A., Torres, M.A., Catena, A., Cándido, A., Maldonado, A., 2018. Electrophysiological brain indices of risk behavior modification induced by contingent feedback. Int. J. Psychophysiol. 124, 43–53. https://doi.org/10.1016/j.ijpsycho.2018.01.003.
Moran, C., Bennett, J.M., Prabhakharan, P., 2019. Road user hazard perception tests: a systematic review of current methodologies. Accid. Anal. Prev. 129, 309–333. https://doi.org/10.1016/j.aap.2019.05.021.
Nævestad, T.-O., Bjørnskau, T., 2012. How can the safety culture perspective be applied to road traffic? Transp. Rev. 32, 139–154. https://doi.org/10.1080/01441647.2011.628131.
Nordfjærn, T., Jørgensen, S., Rundmo, T., 2011. A cross-cultural comparison of road traffic risk perceptions, attitudes towards traffic safety and driver behaviour. J. Risk Res. 14, 657–684.
Norman, D.A., 1988. The Psychology of Everyday Things. Basic Books, New York.
Olson, P.L., Sivak, M., 1986. Perception-response time to unexpected roadway hazards. Hum. Factors 28, 91–96.
Park, D.C., Huang, C.-M., 2010. Culture wires the brain: a cognitive neuroscience perspective. Perspect. Psychol. Sci. 5, 391–400. https://doi.org/10.1177/1745691610374591.
Pelz, D.C., Krupat, E., 1974. Caution profile and driving record of undergraduate males. Accid. Anal. Prev. 6, 45–58. https://doi.org/10.1016/0001-4575(74)90015-3.
Perego, P., Biassoni, F., King, M.J., Ciceri, M.R., 2018. Perception of road hazards in a Tanzanian secondary school before and after a traffic psychology intervention. J. Transp. Health 10, 37–43. https://doi.org/10.1016/j.jth.2018.05.101.
Qu, Y., Lin, L.C., Telzer, E.H., 2019. Culture modulates the neural correlates underlying risky exploration. Front. Hum. Neurosci. 13 https://doi.org/10.3389/fnhum.2019.00171. Article 171.

Reinisch, P., Zahn, P., Schramm, D., 2010. Using a reaction time model for determining a collision avoidance system's brake timing. Ifac Proc. Vol. 43, 477–482. https://doi.org/10.3182/20100712-3-DE-2013.00008.

Savage, S.W., Potter, D.D., Tatler, B.W., 2020. The effects of cognitive distraction on behavioural, oculomotor and electrophysiological metrics during a driving hazard perception task. Accid. Anal. Prev. 138, 105469. https://doi.org/10.1016/j.aap.2020.105469.

Scialfa, C.T., Borkenhagen, D., Lyon, J., Deschênes, M., Horswill, M., Wetton, M., 2012. The effects of driving experience on responses to a static hazard perception test. Accid. Anal. Prev. 45, 547–553. https://doi.org/10.1016/j.aap.2011.09.005.

Scott-Parker, B., Oviedo-Trespalacios, O., 2017. Young driver risky behaviour and predictors of crash risk in Australia, New Zealand and Colombia: Same but different? Accid. Anal. Prev. 99, 30–38. https://doi.org/10.1016/j.aap.2016.11.001.

Serrano, J., Di Stasi, L.L., Megías, A., Catena, A., 2014. Affective-sound effects on driving behaviour. Transport 29. https://doi.org/10.3846/16484142.2013.815133, 100–106–100–106.

Shahar, A., Alberti, C.F., Clarke, D., Crundall, D., 2010. Hazard perception as a function of target location and the field of view. Accid. Anal. Prev. 42, 1577–1584. https://doi.org/10.1016/j.aap.2010.03.016.

Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x.

Shiferaw, B., Downey, L., Crewther, D., 2019. A review of gaze entropy as a measure of visual scanning efficiency. Neurosci. Biobehav. Rev. 96, 353–366. https://doi.org/10.1016/j.neubiorev.2018.12.007.

Shkurko, A., 2020. Mapping cultural values onto the brain: the fragmented landscape. in press Integr. Psychol. Behav. Sci.. https://doi.org/10.1007/s12124-020-09553-0.

Sivak, M., 1996. The information that drivers use: is it indeed 90% visual? Perception 25, 1081–1089. https://doi.org/10.1068/p251081.

Sivak, M., Soler, J., Tränkle, U., 1989a. Cross-cultural differences in driver self-assessment. Accid. Anal. Prev. 21, 371–375. https://doi.org/10.1016/0001-4575(89)90028-6.

Sivak, M., Soler, J., Tränkle, U., 1989b. Cross-cultural differences in driver risk-taking. Accid. Anal. Prev. 21, 363–369. https://doi.org/10.1016/0001-4575(89)90027-4.

Sivak, M., Soler, J., Tränkle, U., Spagnhol, J.M., 1989c. Cross-cultural differences in driver risk-perception. Accid. Anal. Prev. 21, 355–362. https://doi.org/10.1016/0001-4575(89)90026-2.

Stephens, A.N., Ohtsuka, K., 2014. Cognitive biases in aggressive drivers: Does illusion of control drive us off the road? Pers. Individ. Dif. 68, 124–129. https://doi.org/10.1016/j.paid.2014.04.016.

Sullman, M.J.M., Taylor, J.E., 2010. Social desirability and self-reported driving behaviours: Should we be worried? Transp. Res. Part F Traffic Psychol. Behav. 13, 215–221. https://doi.org/10.1016/j.trf.2010.04.004.

Sullman, M.J.M., Hill, T., Stephens, A.N., 2018. Predicting intentions to text and call while driving using the theory of planned behaviour. Transp. Res. Part F Traffic Psychol. Behav. 58, 405–413. https://doi.org/10.1016/j.trf.2018.05.002.

Summala, H., 2000. Brake reaction times and driver behavior analysis. Transp. Hum. Factors 2, 217–226. https://doi.org/10.1207/STHF0203_2.

Svenson, O., Fischhoff, B., MacGregor, D., 1985. Perceived driving safety and seatbelt usage. Accid. Anal. Prev. 17, 119–133.

Titz, J., Scholz, A., Sedlmeier, P., 2018. Comparing eye trackers by correlating their eye-metric data. Behav Res 50, 1853–1863. https://doi.org/10.3758/s13428-017-0954-y.

Torfs, K., Meesmann, U., Van den Berghe, W., Trotta, M., 2016. ESRA 2015–The Results. Synthesis of the Main Findings from the ESRA Survey in 17 Countries. ESRA Project (European Survey of Road Users' Safety Attitudes) (No. 2016- R- 05- EN). Belgian Road Safety Institute, Brussels, Belgium.

Torres, M.A., Megías, A., Catena, A., Cándido, A., Maldonado, A., 2017. Opposite effects of feedback contingency on the process of risky decisions-making. Transp. Res. Part F Traffic Psychol. Behav. 45, 147–156. https://doi.org/10.1016/j.trf.2016.12.007.

Tūskė, V., Šeibokaitė, L., Endriulaitienė, A., Lehtonen, E., 2019. Hazard perception test development for Lithuanian drivers. Iatss Res. 43, 108–113. https://doi.org/10.1016/j.iatssr.2018.10.001.

Underwood, G., Phelps, N., Wright, C., Van Loon, E., Galpin, A., 2005. Eye fixation scanpaths of younger and older drivers in a hazard perception task. Ophthalmic Physiol. Opt. 25, 346–356.

Underwood, G., Crundall, D., Chapman, P., 2011. Driving simulator validation with hazard perception. Transp. Res. Part F Traffic Psychol. Behav. 14, 435–446.

Ventsislavova, P., Crundall, D., Baguley, T., Castro, C., Gugliotta, A., Garcia-Fernandez, P., Zhang, W., Ba, Y., Li, Q., 2019. A comparison of hazard perception and hazard prediction tests across China, Spain and the UK. Accid. Anal. Prev. 122, 268–286. https://doi.org/10.1016/j.aap.2018.10.010.

Williams, J.R., 2008. The Declaration of Helsinki and public health. B World Health Organ 86, 650–652. https://doi.org/10.1590/S0042-96862008000800022.

World Health Organization, 2018. Global Status Report on Road Safety 2018: Summary. World Health Organization, Geneva, Switzerland.

Yamada, Y., Kobayashi, M., 2018. Detecting mental fatigue from eye-tracking data gathered while watching video: evaluation in younger and older adults. Artif. Intell. Med. 91, 39–48. https://doi.org/10.1016/j.artmed.2018.06.005.

Zahreba, V., 2017. What Should Be Public Policy on Road Safety in Ukraine? (Policy Paper). United States Agency for International Development (USAID).