

Cross-cultural effects on drivers' hazard perception

Phui Cheng Lim^a, Elizabeth Sheppard^b & David Crundall^c ¹

^a Corresponding author. Address: Room BB47, Faculty of Science, University of Nottingham Malaysia Campus, Jalan Broga, 43500 Semenyih, Selangor, Malaysia. Tel: +603 8924 8350. E-mail address: khyx1lpn@nottingham.edu.my

^b Address: Room B1A15, Faculty of Science, University of Nottingham Malaysia Campus, Jalan Broga, 43500 Semenyih, Selangor, Malaysia. E-mail address: elizabeth.sheppard@nottingham.edu.my

^c Address: Room 315, School of Psychology, University of Nottingham, University Park, Nottingham NG7 2RD, UK

¹ Present address: Division of Psychology, Chaucer Building, Nottingham Trent University, Nottingham NG1 4BU. E-mail address: david.crundall@ntu.ac.uk

Abstract

Hazard perception tests are used in several developed countries as part of the driver licensing curriculum, however little research has been done in developing countries where road safety is a primary concern. We conducted a cross-cultural hazard perception study to examine the transferability of hazard perception skills between Malaysia and the UK, using hazard clips filmed in both countries. The results showed that familiarity with both the driving environment and type of hazard facilitated drivers' ability to discriminate hazards in a timely manner, although overall drivers viewed and responded to hazards largely similarly regardless of origin. Visual strategies also appeared to be moderated mainly by the immediate driving environment rather than driver origin. Finally, Malaysian drivers required a higher threshold of danger than UK drivers before they would identify a situation as hazardous, possibly reflecting the more hazardous road environment in Malaysia. We suggest that hazard perception testing in developing countries requires a test where performance cannot be confounded with differing thresholds for hazardousness.

Keywords: hazard perception, driving experience, cross-cultural, hazard familiarity, environmental familiarity, eye movements

Cross-cultural effects on drivers' hazard perception

1. Introduction

The hazard perception (HP) skill of a driver refers to the ability to identify potentially dangerous situations on the road. It is typically assessed by showing participants video clips of hazards, and asking them to respond as soon as they detect a developing hazard, with shorter response times reflecting greater levels of HP skill (Chapman & Underwood, 1998; Horswill & McKenna, 2004; McKenna, Horswill, & Alexander, 2006; Sagberg & Bjørnskau, 2006; Wetton, Hill, & Horswill, 2011). According to Horswill and McKenna (2004), hazard perception is one of the only components of driving skill that has been consistently linked to accident involvement across multiple studies (Horswill, Anstey, Hatherly, & Wood, 2010; McKenna & Horswill, 1999; Quimby, Maycock, Carter, Dixon, & Wall, 1986). A prospective study by Drummond (2000) also found that newly licensed drivers' hazard perception performance was linked to their likelihood of being involved in a fatal collision in the subsequent 12 months. Additionally, past research has found that experienced drivers outperform novices in hazard perception tests (Horswill et al., 2008; Renge, 1998; Wallis & Horswill, 2007), with driving experience considered as a proxy for driver safety, although other studies have failed to find this experiential difference (Chapman & Underwood, 1998; Crundall, Underwood, & Chapman, 2002; Sagberg & Bjørnskau, 2006). It has recently been suggested that the lack of replication may stem in part from the particular hazards used in these studies, with certain types of hazard differentiating experience more effectively than others (Borowsky, Shinar, & Oron-Gilad, 2007; Crundall et al., 2012). Nevertheless, the reported predictive utility of hazard perception testing has led to the inclusion of a hazard perception component in driver licensing in the UK, Australia and the

Netherlands, where there has been some evidence to suggest its efficacy in reducing accidents in new drivers (Wells, Tong, Sexton, Grayson, & Jones, 2008).

Notably, the research cited has been conducted in developed countries where road safety is relatively mature, yet the vast majority of road fatalities worldwide occur in developing countries (Nantulya & Reich, 2002; Peden et al., 2004; Toroyan, 2009). While some cross-cultural research on driving has been conducted comparing developed and developing countries (Lund & Rundmo, 2009; Nordfjærn & Rundmo, 2009; Özkan, Lajunen, Chliaoutakis, Parker, & Summala, 2006; Sivak, Soler, Tränkle, & Spagnhol, 1989), this has primarily examined risk-taking behavior and/or self-reported perceptions of driving skill and traffic risk. Hazard perception has yet to be investigated cross-culturally in depth, even within developed countries, and little is known about the transferability of hazard perception skills between noticeably different countries and cultures; are there crucial underlying skills that can successfully transfer between countries, or are strategies and skills culturally distinct?

Exploring hazard perception cross-culturally can also shed light on how location familiarity impacts on hazard perception skill. Wetton et al. (2010) found novice/experienced latency differences in Australian participants when using footage of both Australian and UK roads, suggesting the advantage of experience endures even in unfamiliar environments and hazard perception abilities contain at least some general component. However, several questions remain unanswered. For instance, the UK and Australian settings used by Wetton et al. (2010) are very similar. Cultures, road laws, vehicles, driving styles, and even architecture overlap considerably between the two countries. Would similar transference of skill occur in vastly differently settings? If a locational advantage does exist, is it due to familiarity with the driving environment itself (which we will term *environmental* familiarity), familiarity with hazards

typically encountered in that environment (*hazard familiarity*) or, likely, some combination of the two?

The current study aims to investigate the extent to which hazard perception skills transfer cross-culturally, using the typical reaction time paradigm, across two highly different settings: the UK and Malaysia. This will reveal similarities and differences in hazard perception performance between drivers from two very different driving cultures, and will hopefully identify whether these are specific to the context (i.e. can UK drivers perform well on both UK and Malaysian HP clips?).

Hazard perception performance of experienced and novice drivers from both the UK and Malaysia will be compared using the reaction time paradigm described earlier with video footage from both the UK and Malaysia. As a former British colony, Malaysia shares several commonalities with the UK that make it a suitable comparison point; namely, similar road rules and left-hand driving environment. It is also a middle-income country with a high percentage of car ownership. However, Malaysia has drastically higher accident rates; in 2010 its road fatality rate was 24 per 100,000 people (Sarani, Syed Mohamed Rahim, Mohd Marjan, & Wong, 2012), while the UK's was 3 per 100,000 (Kilbey, 2011), suggesting many more dangerous events and a generally more hazardous road environment. While many factors likely contribute to this discrepancy, given an eightfold difference in fatality rates we expect to see at least some difference in hazard perception skills between the populations. By comparing Malaysia and UK drivers' hazard perception abilities in Malaysian and UK road environments, we should obtain further insight into hazard perception transferability across cultures. Wetton et al. (2010)'s findings certainly suggest some amount of transferability, although this was seen in the UK and Australia where accident rates are very similar. In a more hazardous environment, hazard

perception skills are arguably even more critical, although it is also possible that in more disparate driving environments, location familiarity may play a bigger role than Wetton et al. (2010) found.

While we expect to see the typical effect of experience in both locations, similar to Wetton et al. (2010), we also expect location familiarity to confer a significant advantage in these two distinct cultures, and we therefore hypothesize superior performance in the form of shorter reaction times and higher response rates when participants view clips from their home country. Furthermore, while Malaysian drivers certainly experience more hazards than UK drivers, this effect could play out in either direction: they may be quicker to detect hazards due to their greater exposure to them, or equally, they may be desensitized to hazards and have a higher criterion for hazard identification compared to UK drivers.

Additionally, we can break down pure location familiarity and infer its subcategories of environmental and hazard familiarity, as mentioned earlier. Half of the clips from both the UK and Malaysia were matched for hazard content. For example, in one matched pair of clips, a car on the highway overtakes the camera car on the inside lane. The other half of the clips consisted of unmatched hazards that are more representative of the different countries. For instance, motorcycle and scooter riders are relatively infrequent hazards in the UK, while in Malaysia powered two-wheelers make up a significantly greater proportion of the traffic. Similarly, zebra crossings are relatively common in the UK, but extremely rare in Malaysia compared to pelican crossings. We expect both environmental and hazard familiarity to confer an advantage, and hypothesize that all drivers will exhibit superior performance when viewing matched hazards in their home country, compared to matched hazards in their non-home country (environmental familiarity), and also when viewing matched non-home country hazards compared to unmatched

non-home country hazards (hazard familiarity). Furthermore, we hypothesize the greatest performance difference between Malaysian and UK drivers when viewing unmatched hazards, as these presumably confer both environmental and hazard familiarity.

Finally, the current study also investigates visual strategies by using eye tracking measures for all participants. This provides a greater insight into hazard perception skill than a measure of response time can provide. For instance, we expect participants to have shorter fixation durations in their home environment, indicating greater processing efficiency (Crundall & Underwood, 1998). We also expect to see the typical effects of attentional capture across all clips, i.e. longer fixation durations upon hazard onset, although this should be less pronounced in experienced drivers (Chapman & Underwood, 1998). It may also help establish whether differences in hazard perception performance stem from differences underlying attentional processes (how quickly participants spot the hazard) or differences in criterion (how quickly participants judge the hazard to pose a threat).

2. Methods

2.1. Participants

Forty-five participants were recruited from the UK and 55 from Malaysia, all of whom held either full or provisional driving licenses from their respective countries and had normal or corrected-to-normal vision. Participants were split into two further sub-groups consisting of novice and experienced drivers, resulting in four groups in total: UK novice drivers ($n = 20$, mean age of 18.9 years and average licensing time of 9.1 months), UK experienced drivers ($n = 25$, mean age of 21.0 years and average licensing time of 45.6 months), Malaysia novice drivers ($n = 26$, mean age of 18.8 years and average licensing time of 9.8 months) and Malaysia experienced drivers ($n = 27$, mean age of 21.9 years and average licensing time of 49.4 months).

2.2. Stimuli

A Panasonic HD SDC-600/900 camera was attached to the windscreen of various cars during journeys made around the UK and Malaysia. Footage was shot in the daytime, under clear weather conditions and normal visibility. Twenty clips from Malaysia and 20 clips from the UK, each containing one hazard and ranging from 6 – 54 seconds in length, were selected and edited from the resulting footage. Table 1 contains clip length information for all clip categories.

Of the 20 clips from each country, 10 were matched for hazard content and what happened in the clip. For example, in one matched pair of clips, a pedestrian steps out from behind a parked bus directly in front of the camera car. The other 10 clips from each country were not matched and reflected hazardous situations that were selected without restriction, and were subsequently more representative of the different countries. For instance, in an unmatched Malaysia clip an oncoming motorcyclist pulls out from behind a truck and into the path of the camera car, while in an unmatched UK clip a pedestrian with a bicycle steps into the road at a zebra crossing, forcing the camera car to stop and allow the pedestrian to pass. Matched clip pairs and unmatched clips were selected independently by one Malaysian researcher and one UK researcher, both of whom held a full driving license in their respective countries. Disagreements were resolved and the final clips, both matched and unmatched, selected via discussion between the above two researchers and a further two who had spent a significant amount of time driving in both countries.

2.3. Apparatus

The stimuli were played on a 17" TFT monitor at a resolution of 1024 x 768, presented using Tobii Studio 2.3. Participants were seated 65 cm from the screen at a visual angle of 29.6°

x 23.2 °, and their eye movements were monitored using a Tobii T60 eye tracker in Malaysia and Tobii T1750 eye tracker in the UK, which sample at 60 Hz and 50 Hz respectively. While this results in a slightly higher margin of error for UK participants (20 ms compared to 16.67 ms), this should not introduce any systematic bias into the results.

2.4. Design

A 2 x 2 x 2 x 2 mixed design was used. The between-groups factors were the country of origin of the driver (Malaysia or UK), and experience level (novice or experienced). The within-groups factors were the country where the clip was filmed (Malaysia or UK) and whether the clip was matched or unmatched.

The stimuli were separated by country into two blocks of 20 clips. Within each block, the order of clips to be presented was randomized using a Latin square. The order of the blocks was counterbalanced across participants.

A *hazard window* was defined for each hazard, similar to the “danger window” used by Chapman and Underwood (1998). A single researcher defined all forty windows, of which a sub-sample were rated by a second researcher with a high degree of agreement. The window began at the earliest point in time where the hazard was detectable by the viewer and clearly on a course that would eventually obstruct the camera car. For instance, a pedestrian walking beside the road would not be considered a hazard until he steps towards the road, or makes any other movement that sets him on a collision course with the camera car. The window ended at the point at which a braking or avoidance response by the driver would no longer prevent a collision. Table 1 depicts the average length of each hazard window for each clip category.

Four behavioral measures were analyzed: responses to pre-defined hazards (responses were made within the hazard window and the participant verbally reported the appropriate

hazard at the end of the clip), extra-hazard response rates (number of button presses made that were not in response to the pre-defined hazard), reaction time and number of extra-hazard responses for every pre-defined hazard response. Two eye tracking measures were analyzed: time for participants to first fixate the hazard (regardless of whether they identified it as one), and the average fixation duration prior to and directly after hazard onset. Hazard precursors were not included in the analyses.

2.5. Procedure

After giving informed consent, participants were seated in front of the eye tracker and instructed to watch the video clips and click the left mouse button as soon as they saw a hazard developing. Participants were told to respond to “hazardous events, i.e. situations in which the driver may need to change speed or direction of their vehicle to prevent a potential collision,” similar to the definition utilized by Wetton et al. (2011), although modified to allow for lesser English proficiency. They were informed that at the end of each video, they would be asked to briefly describe the last hazard they had responded to.

Commented [LS1]: Should probably specify what's the difference

Participants were then calibrated on the eye tracker, using a standard 9-point calibration procedure. Before starting each 20-clip block, participants were shown two sample clips for them to practice the task, both taken from the same country as the block they were about to view. These practice clips were not used in the subsequent experiment and were the same for all participants. After the practice trials were completed, participants began the first block of clips. Participants were re-calibrated on the eye tracker in-between blocks.

3. Results

Means and standard deviations for all measures are summarized for novice drivers in Table 2 and experienced drivers in Table 3. Unless otherwise stated, a 2 (driver origin:

Malaysia or UK) x 2 (experience: novice or experienced) x 2 (country clip: Malaysia or UK) x 2 (matching: matched or unmatched) mixed ANOVA was run for all measures.

3.1. Behavioral analysis

3.1.1. Pre-defined hazard responses

An analysis of responses to pre-defined hazards (where a correct answer is indicated by a button response during the hazard window, and a correct verbal identification following the clip), identified two main effects. First, UK drivers identified more hazards than Malaysian drivers ($F_{1,94} = 47.04, p < .001, \eta^2_p = .334$) and secondly, all drivers identified more matched hazards than unmatched ($F_{1,94} = 64.87, p < .001, \eta^2_p = .408$). There were two two-way interactions, depicted in Figure 1. Panel (a) shows an interaction between matching and clip country ($F_{1,94} = 58.43, p < .001, \eta^2_p = .383$), where participants responded equally often to both matched and unmatched Malaysian hazards, but more often to UK matched hazards than UK unmatched ($t(97) = 10.93, p < .001, d = .823$). A second interaction, shown in panel (b), was found between driver origin and clip country ($F_{1,94} = 12.39, p = .001, \eta^2_p = .116$), where participants responded to hazards filmed in their home country more often. This effect was significant in participants from both countries ($t(52) = 2.81, p = .007, d = .278$ for Malaysia drivers and $t(44) = -2.19, p = .034, d = .392$ for UK drivers).

Previous studies have however raised problems with using a restrictive window for calculating responses to hazards, as some exceptional individuals may be penalized for responding too soon, while some drivers may recognize the hazard but forget to respond (Jackson, Chapman, & Crundall, 2009). A second analysis was therefore undertaken which included all trials where participants correctly identified the pre-defined hazard in their later verbal description, regardless of when or whether they had made a button press response during

the clip. This, more liberal, analysis was chosen to ensure that response time outliers did not confound the response rate measure.

Results are depicted in Figure 2. As experience had no effect, means are shown collapsed across experienced and novice groups.

As in the previous analysis, main effects of driver origin ($F_{1,94} = 39.33, p < .001, \eta^2_p = .295$) and matching ($F_{1,94} = 112.47, p < .001, \eta^2_p = .545$) were found, where UK drivers identified more hazards, and there were more responses to matched hazards. There was also a main effect of clip country ($F_{1,94} = 13.31, p < .001, \eta^2_p = .124$), where drivers identified Malaysian hazards more often. An interaction between matching and clip country was again found ($F_{1,94} = 71.51, p < .001, \eta^2_p = .432$), showing the same pattern as this interaction in the previous analysis. This was subsumed by a three-way interaction between matching, clip country and driver origin ($F_{1,94} = 4.17, p = 0.044, \eta^2_p = .042$), where all drivers responded to UK unmatched clips the least compared to the other three clip groups. While this was significant for both groups of drivers when compared to performance on unmatched clips filmed in Malaysia ($t(52) = 6.41, p < .001, d = .763$ for Malaysia drivers and $t(44) = 4.41, p < .001, d = .768$ for UK drivers), the drop in response rate is numerically greatest in Malaysian drivers.

3.1.2. Number of extra-hazard responses for every response to pre-defined hazards

The extra-hazard response rate was the number of button press responses made during an entire video that were not a response to the pre-defined hazard. The number of extra-hazard responses made for every pre-defined hazard response, using the initial strict measure of response rate, was calculated for each participant. A higher number therefore means more extra-hazard responses made relative to pre-defined hazard responses.

Three main effects were found: matching, where unmatched clips had a higher rate of responses ($F_{1,94} = 25.15, p < .001, \eta^2_p = .211$), country clip, where Malaysia clips had a higher rate of responses ($F_{1,94} = 6.05, p = .016, \eta^2_p = .060$) and driver origin, where UK drivers had a higher rate of responses ($F_{1,94} = 8.76, p = .004, \eta^2_p = .085$). Two two-way interactions were found, country clip and driver origin ($F_{1,94} = 8.61, p = .004, \eta^2_p = .084$) and matching and country clip ($F_{1,94} = 27.19, p < .001, \eta^2_p = .224$). These interactions were subsumed by a three-way interaction depicted in [Figure 3](#) of country clip, matching and driver origin ($F_{1,94} < .001, \eta^2_p = .138$), where for matched clips, all drivers made more extra-hazard responses in UK clips and UK drivers made more extra-hazard responses than Malaysian drivers; however, for unmatched clips, UK drivers watching Malaysian clips had a particularly high extra-hazard response rate ($t(44) = 5.93, p < .001, d = .904$ compared to UK unmatched clips), but there was no difference between Malaysian drivers watching UK and Malaysian clips, or UK drivers watching UK clips.

An additional three-way interaction of driver origin, driver experience level, and matching is depicted in [Figure 4](#) ($F_{1,94} = 8.37, p = .005, \eta^2_p = .082$), where novice experienced UK drivers had a higher response rate on unmatched clips compared to unmatched experienced Malaysian drivers and UK novices showed no such difference ($t(25) = -4.23, p < .001, d = .719$ and $t(24) = -6.34, p < .001, d = .335$ respectively), but experienced Malaysian drivers and UK novices showed no such difference.

3.1.3. Reaction time

Reaction times were calculated from the beginning of the hazard window, based on the initial strict measure of response rate. Only clips that had responses to the pre-defined hazards during the hazard window were included; all other responses were removed, as were clips with no responses.

Two main effects were found: matching ($F_{1,94} = 20.44, p < .001, \eta^2_p = .179$), where drivers responded to matched hazards faster, and driver origin ($F_{1,94} = 53.94, p < .001, \eta^2_p = .365$), with UK drivers having faster overall response times. There was a crossover interaction of clip country and matching ($F_{1,94} = 49.63, p < .001, \eta^2_p = .346$), where drivers reacted fastest to Malaysian matched clips out of all clips ($t(97) = -5.98, p < .001, d = .567$ compared to UK matched clips) and slowest to Malaysia unmatched clips ($t(97) = 3.67, p < .001, d = 1.13$ compared to UK unmatched clips) but there was no difference in reaction time for UK clips regardless of matching.

3.2. Eye movement analysis

Eye movement analysis was conducted using Tobii Studio's Dynamic Area of Interest (AOI) tool. 30 participants were removed due to missing eye tracking data in several videos, leaving 68 participants in total (16 UK novices; 19 UK experienced; 13 Malaysia novices; 20 Malaysia experienced). All behavioral analyses were re-run with these participants to confirm that the results followed the same pattern as above.

Two analyses were conducted on the eye tracking data: time taken to first fixate the pre-defined hazards and the average fixation duration prior to and directly after hazard onset.

3.2.1. Time to first fixate on pre-defined hazards

The time taken to first fixate on pre-defined hazards was measured from hazard onset, until the participant fixated on the hazard for the first time. If there were no fixations on the hazard during the hazard window, that particular trial was removed.

Two main effects and a subsequent interaction were found. Matched hazards ($F_{1,64} = 18.21, p < .001, \eta^2_p = .221$) were fixated faster, as were Malaysian hazards ($F_{1,64} = 10.31, p = .002, \eta^2_p = .139$). This was mainly driven by an interaction of matching and clip country ($F_{1,64} =$

11.11, $p = .001$, $\eta^2_p = .148$), where Malaysian matched hazards were fixated particularly quickly compared to both Malaysian unmatched hazards ($t(67) = -6.68$, $p < .001$, $d = .730$) and UK matched hazards ($t(67) = -4.46$, $p < .001$, $d = .681$), but there was no difference between time to first fixation for all other clip types.

3.2.2. Average fixation duration

Participants' average fixation duration was calculated for two windows during each clip: first from the start of the clip until hazard onset (*pre-onset*), and secondly from hazard onset to offset, i.e. the hazard window described earlier (*post-onset*), as in Chapman and Underwood (1998). For this particular measure, the ANOVA used in all previous analyses was conducted but with the additional within-groups factor of pre- or post- hazard onset, resulting in a 5-way mixed ANOVA.

A main effect of clip country was found, with shorter average fixations during Malaysian videos ($F_{1,64} = 8.97$, $p = .004$, $\eta^2_p = .123$). Four two-way interactions were found, two shown in [Figure 5](#). Panel (a) depicts an interaction of clip country and driver origin ($F_{1,64} = .021$, $\eta^2_p = .081$), where Malaysian drivers had particularly short fixations in Malaysian clips, compared to UK clips ($t(32) = -4.43$, $p < .001$, $d = .412$). There was also a crossover interaction of experience and pre-/post- hazard onset ($F_{1,64} = 4.71$, $p = .034$, $\eta^2_p = .069$), shown in panel (b), where novice drivers had longer fixations pre-onset, but experienced drivers showed the opposite pattern; however, post-hoc tests revealed no significant differences. The remaining two-way interactions, pre-/post-onset and country clip ($F_{1,64} = 10.60$, $p = .002$, $\eta^2_p = .142$) and matching and country clip ($F_{1,64} = 24.29$, $p < .001$, $\eta^2_p = .275$), were subsumed by a three-way interaction of pre-/post- onset, matching and clip country ($F_{1,64} = 18.82$, $p < .001$, $\eta^2_p = .227$), shown in [Figure 6](#). There were no pre- or post- onset differences in Malaysia matched or UK matched

clips, but participants showed shorter fixations post-onset than pre-onset in Malaysia unmatched clips ($t(67) = 2.25, p = .028, d = .154$) and longer fixations post-onset than pre-onset in UK unmatched clips ($t(67) = -3.91, p < .001, d = .295$).

4. Discussion

In this discussion, we will focus on how experience and familiarity with both driving environment and hazard types affects drivers' performance. We will also review factors that appear to moderate visual strategies, and the possibility of "look but fail to see" errors (Brown, 2002). Finally, we will discuss how different hazard identification thresholds may affect participants' responses, and the subsequent implications for hazard perception testing in developing countries.

4.1. Hazard perception skill transferability and the effect of familiarity

In line with Wetton et al. (2010)'s findings, hazard perception skill appears to be highly transferable, as Malaysian and UK drivers showed the same general response pattern for most measures. Differences between drivers were mostly seen when one country's drivers had an exaggerated response to particular clip types without changing the general response pattern; for instance, UK unmatched hazards produced the lowest response rates across all participants, but Malaysian drivers had particularly low response rates on these clips. Performance did vary with all independent variables, however within-groups factors (i.e. clip country and matched/unmatched clips) accounted for more variation than between-groups factors (driver origin and experience), suggesting that while certain nuances of hazard perception are affected by context, drivers view and respond to hazards largely similarly regardless of their home driving environment.

Drivers detected more pre-defined hazards from their own country (although only in the stricter, time-bound analysis) and also hazards that were matched for type. One possible explanation for this is that both environmental and hazard familiarity facilitate drivers' ability to detect hazards in a timely manner. It is also possible that drivers have a lower threshold for identifying hazards when the hazards and/or environment are familiar. However, if this were the case we would also expect to see more extra-hazard responses in home countries. In fact, all participants made more responses, to both pre-defined and extra-hazards, in Malaysian clips, and this tendency was especially pronounced in UK drivers, possibly reflecting a greater general inclination for UK drivers to appraise events as hazardous, and/or the more hazardous road environment in Malaysia. If, then, we rule out the possibility that drivers have lower thresholds for identifying hazards in a familiar driving environment, we can consider that environmental familiarity facilitates drivers' ability to discriminate between hazards and non-hazards effectively. While Malaysian drivers appear to discriminate equally well in both countries, UK drivers do appear to discriminate more effectively in UK clips. Furthermore, the interaction between driver origin and clip country is only seen in the stricter, time-bound analysis and not the more liberal analysis; given the ability to retrospectively identify hazards with no time pressure, the advantage of environmental familiarity disappears (in fact, it disappears altogether for UK drivers, who detect more pre-defined Malaysian hazards than UK hazards in the more liberal analysis). This suggests that environmental familiarity does indeed prime drivers to react appropriately once a hazard appears, possibly because their mental models are richer in familiar environments; having a greater awareness of the possible hazards and precursors that may occur makes them better equipped to detect early warning signs of dangerous situations (Underwood, Chapman, Bowden, & Crundall, 2002).

It should be noted however that the driver origin / clip country interaction does *not* occur in reaction time analyses, so environmental familiarity does not necessarily mean that drivers react faster, simply that they react within an appropriate timeframe. Perhaps the interaction does not stem from the advantage of a familiar environment, but rather, the disadvantage of an unfamiliar one; it is possible that processing an unfamiliar environment increases cognitive load to the point where drivers may occasionally (but not consistently) fail to discriminate hazards in time.

Hazard familiarity may also affect hazard identification thresholds, although evidence for this is inconclusive. Participants identified less unmatched hazards; this was largely driven by UK unmatched clips, which Malaysian drivers performed particularly poorly on, suggesting that they were less likely to appraise hazards as hazardous when they were unaccustomed to the type of event occurring. Drivers may be more inclined to react to familiar hazards because they are aware of the possible dangers these hazards pose and conversely, less inclined to identify unfamiliar hazards because they lack the necessary experience to appraise them as hazards (Groeger, 2000).

4.2. Experience

Contrary to our hypothesis, it appears that driving experience did not have a significant effect on response latency. However, experience did play a role in several interactions, suggesting it may be important cross-culturally but only within specific hazard contexts; for instance, UK experienced drivers appear to be especially sensitive to hazards when both location and hazard type are unfamiliar.

Finding no effect of experience in response rate and reaction time might be explained in Malaysian drivers by the relatively hazardous Malaysian driving environment, compared to the

developed countries where previous hazard perception research has been conducted. It is somewhat surprising that there also was no effect of experience among drivers in the UK, where the reaction time paradigm has been found to differentiate experience in the UK and other developed countries (Horswill et al., 2008; Wallis & Horswill, 2007; Wetton et al., 2010). This may be due to an overall lack of experience, as the experienced driver group in this study had an average post-license experience of approximately four years, compared to the above studies where the experienced groups had held their license for over ten years. Indeed, other studies in developed countries have also failed to find significant differences between experienced and novice drivers when the experienced group had held their license for a relatively short period of time (Chapman & Underwood, 1998; Crundall et al., 2002).

However, yet other studies have also failed to find a difference of experience even with a highly experienced group (Crundall, Chapman, Phelps, & Underwood, 2003; Sagberg & Bjørnskau, 2006), and it is also possible that the nature of the hazards used may explain the failure to replicate differences of experience. Crundall et al. (2012) found that certain hazard types differentiated experience more successfully than others, reporting that behavioral prediction (BP) hazards, where the hazard's precursor and hazard itself are the same stimulus, were the least effective. Conversely, environmental prediction (EP) hazards, where the precursor and hazard are indirectly related, were detected faster and more often with greater levels of driving experience. In this study, the majority of clips contained BP hazards; out of forty hazards, only five were EP. It is therefore possible that, among other factors, the predominance of BP hazards contributed to the lack of differentiation between novice and experienced drivers.

4.3. Finally, it should be noted that all UK participants had passed a hazard perception test in order to obtain their driving license. UK novice participants in particular would have

Commented [LS2]: This point seems to complement the previous one, rather than be a however.

Commented [PCL3]: For own reference:
 Chapman & Underwood 1998: Relatively inexperienced experience group
 Crundall et al 2003: Police vs experienced vs novice, all had same ratings of hazardness (measured throughout using a sliding scale)
 Crundall et al 2002: Exp drivers 56 months, relatively low
 S & B 2006: 27.1 years experienced group, trended only

practiced for, and passed, a hazard perception test less than one year prior to taking part in the study; hence practice effects may have facilitated their performance. This raises the interesting question of whether is still feasible to validate hazard perception tests in the UK on the basis of driving experience due to the pervasive nature of hazard perception testing over the last decade. Visual strategies

Previous research has found that visual strategies vary by both experience and driving environment (Borowsky, Shinar, & Oron-Gilad, 2010; Chapman & Underwood, 1998; Crundall, Chapman, France, Underwood, & Phelps, 2005; Crundall & Underwood, 1998; Konstantopoulos, Chapman, & Crundall, 2010; Underwood et al., 2002; Underwood, Chapman, Brocklehurst, Underwood, & Crundall, 2003) and thus we hypothesized differences for both experience level and country clip. However, eye tracking metrics did not differ substantially between drivers, and varied mostly between clip country and whether the clips were matched. Again, it is surprising to find no difference in eye movement patterns between novice and experienced drivers, as the studies cited above all found differences in visual strategies. However, as mentioned earlier it is possible that the experienced drivers in this study had not been driving long enough to have sufficiently developed visual strategies; in all studies cited above the experienced group of drivers had held their license for at least 4.5 years longer than the novice group, often substantially more. Notably, Konstantopoulos, Chapman and Crundall (2012) failed to find differences between novice and experienced drivers with a licensing time difference of 4 years, although this may have been due to a floor effect of the task.

Fixations were shorter in Malaysian videos, although this was largely driven by Malaysian drivers having particularly short fixations in these clips; their eye movements may have been more efficient in a familiar environment, allowing them to process more visual

Commented [PCL4]: MY drivers had particularly short fixations in MY clips and particularly long in UK clips, perhaps showing LESS adaption to the environment?

information in a shorter time (Borowsky, Oron-Gilad, Meir, & Parmet, 2011; Chapman & Underwood, 1998; Crundall et al., 2012; Konstantopoulos et al., 2010). As the UK road environment is less cluttered than Malaysia's and therefore less visually demanding, it is possible that Malaysian drivers adapted their visual strategies to the current environment, similar to the behavior Crundall and Underwood (1998) found in experienced drivers. Interestingly, both novice and experienced UK drivers failed to show this effect and had similar length fixations across both countries' clips.

Participants also had longer fixations upon hazard onset in UK unmatched clips, showing the typical effect of attentional capture (Chapman & Underwood, 1998). Unexpectedly, this was the only clip type where this effect was found; there was no difference in pre- and post- hazard onset in Malaysia and UK matched clips, and in fact the opposite effect was found in Malaysia unmatched clips where fixations actually became *shorter* upon hazard onset. Furthermore, novice drivers also had shorter fixations upon hazard onset. Both these results are striking as the effect of attentional capture has been consistently found across several studies (Chapman, Underwood, & Roberts, 2002; Chapman & Underwood, 1998; Underwood, Phelps, Wright, Van Loon, & Galpin, 2005), with longer fixations during dangerous situations, and novice drivers are particularly susceptible to this effect.

It is possible that fixation length did not increase upon hazard onset in Malaysia because of the nature of the driving environment, which often demands attention be divided among several hazards and possible hazards; hence defining a "danger window" as Chapman and Underwood (1998) did may be redundant, as large parts of the Malaysian videos could be considered as hazardous as the original danger window. The frequency of extra-hazard responses to Malaysian unmatched clips certainly seem to suggest as much. In the UK in

contrast, hazards would have been viewed as relatively more hazardous compared to the driving environment, and hence we see a normal effect of attentional capture where fixations increase with hazard onset. However, this does not explain why UK matched videos do not show similar attentional focusing, nor why fixations actually decrease upon hazard onset in the Malaysia matched clips. Further research is required to explore this effect in more depth.

Notably, there were no main effects of driver origin in any of the eye tracking metrics, a sharp contrast from the behavioral results where this effect was found in every metric. This result is interesting in itself, as it suggests that visual strategies are largely moderated by the immediate driving environment, and less so the familiarity of the environment; although both clearly interact. As mentioned, it is possible that the participants in this study had not been driving long enough to have sufficiently developed visual strategies; if this is the case, there are two possible implications. Firstly, they may not have yet developed strategies specific to their environment (assuming they would at all), and would therefore have more flexible search patterns. Secondly, they might use the same search strategy across all environments (Crundall & Underwood, 1998; Falkmer & Gregersen, 2005), although this seems unlikely as Malaysian drivers did appear to adapt their visual strategies to the UK environment. Regardless, drivers' early visual strategies appear to be similar regardless of where they learned to drive. Furthermore, the lack of driver origin differences implies that Malaysian and UK drivers are using broadly similar visual strategies, and therefore these cannot explain the behavioral differences in Malaysian and UK drivers.

4.4. Hazard perception ability and hazard appraisal

Results for response rate and reaction time certainly suggest that UK drivers possess superior hazard perception abilities to Malaysia drivers, as they detected pre-defined hazards in

both countries faster and more often. However, other measures such as extra-hazard responses and eye tracking data suggest that different hazard identification thresholds could also be a major contributing factor to performance. This also indicates cross-cultural differences in how dangerous Malaysian and UK drivers perceive hazards to be.

Overall, Malaysian participants appear to have a substantially higher threshold for hazard identification than UK participants: they had many more non-responses to hazards than UK drivers and a lower extra-hazard response rate, suggesting that an event (regardless of whether it was actually a hazard) needed to reach a higher level of hazardousness before Malaysian drivers would be willing to identify it as a hazard. However, this also resulted in superior performance in discriminating between hazards and non-hazards, as Malaysian drivers made less extra-hazard responses for each pre-defined hazard response, compared to UK drivers. These tendencies are likely due to a greater number of hazards in Malaysia and a more hazardous driving environment overall. Wallis and Horswill (2007) also found evidence for a similar response criterion bias affecting hazard perception performance, although they compared novice and experienced drivers rather than drivers of different cultures. If Malaysian drivers do indeed have a higher threshold for hazard identification this could explain at least in part their lower performance in both response rate and reaction times, as hazardous situations would need to progress further before Malaysia drivers would react, if indeed they did at all.

It is also possible that UK and Malaysian drivers have an altogether different concept of what constitutes a hazard, making the difference one of categorization rather than acceptable thresholds. Indeed, it is debatable as to whether the performance exhibited by UK drivers truly represents superior hazard perception ability, a lower threshold for hazard identification, an altogether different categorization of hazards, or a tendency to over-respond to hazards; clearly,

further research is required to separate these effects. However, in practice, a lowered response criterion and tendency to over-respond presumably stem from a greater degree of cautiousness on the road, which results in safer driving overall. The difficulty of whether HP performance reflects actual skill or different criterion certainly has implications for future HP tests (see section 4.6); however superior HP performance likely has much the same effect, regardless of its cause, in actual on-road situations.

4.5. "Look but fail to see" in a hazard perception context

An eye tracking analysis also found that when drivers failed to identify a pre-defined hazard, in 96.5% of these cases they still fixated on the hazard during its hazard window at least once: so the vast majority of these non-responses were not due to a failure to visually detect the event, but rather a failure to perceive it as hazardous. We should however distinguish between appraising events as non-hazardous and failing to process them altogether, i.e. a "look but fail to see" error (Brown, 2002). While the latter seems unlikely as the hazards were the most salient events in the clip and participants had only a single task, analysis of fixation length also failed to find the effect of attentional focusing described by Chapman and Underwood (1998) for three of the four clip types, suggesting that drivers may have not recognized the hazardousness of the events in the clips. A further study found that fixating on objects in a driving environment did not necessarily facilitate recall and presumably, processing of them (Underwood, Chapman, Berger, & Crundall, 2003). Evidence therefore supports both possibilities.

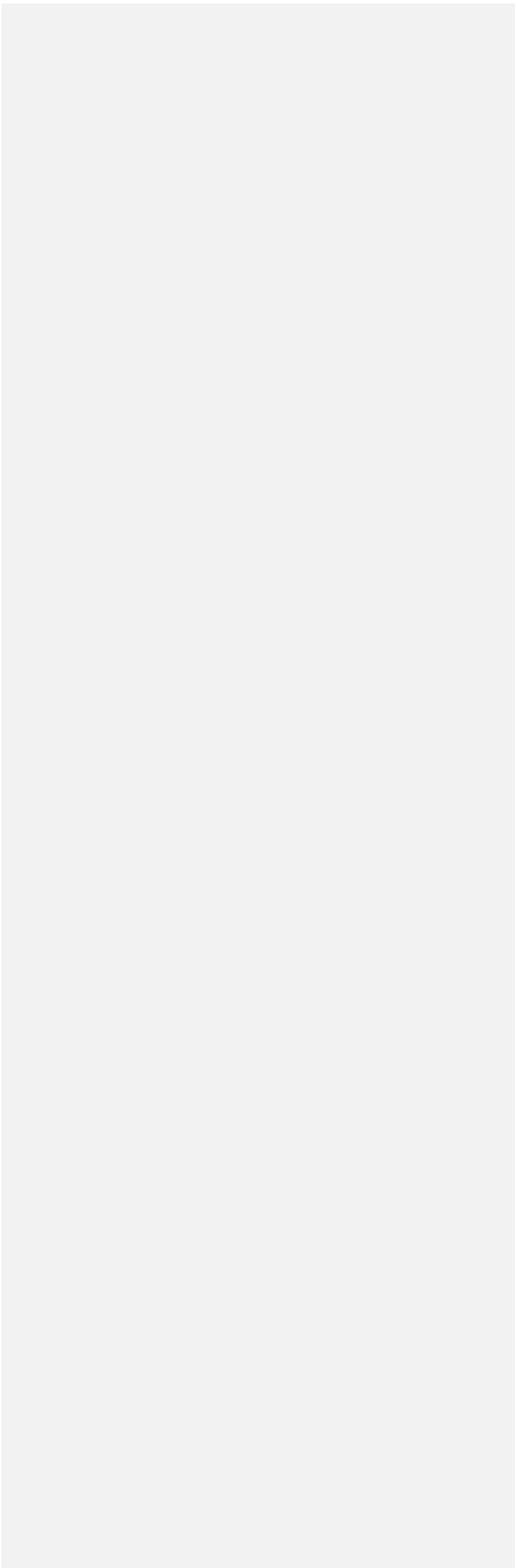
4.6. Hazard perception: A possible diagnostic tool in Malaysia?

The higher hazard criterion among Malaysia drivers poses a challenge to developing a test that differentiates between experienced and novice drivers, as it is possible that this criterion may supersede any differences of experience that might otherwise be found. Within the present

study it is difficult if not impossible to differentiate actual hazard perception ability from hazard identification thresholds, raising validity issues should the test in its current form be used in Malaysia.

This has wider implications for the export of hazard perception methodologies. If drivers in developing countries with more hazardous road environments appraise hazards similarly to Malaysia drivers (i.e. have higher thresholds for what constitutes a hazard), this tendency will confound hazard perception latencies when applying the reaction time paradigm cross-culturally. It does appear that experience can be differentiated among Malaysia drivers, as experience did play a part in certain measures; however, while the issue of differing hazard thresholds remains, a reaction time paradigm may not be suitable as a diagnostic tool in Malaysia or other countries with similarly hazardous driving environments.

It is clear that a test is needed where performance cannot be confounded by hazard identification thresholds. Wetton et al. (2010) have suggested breaking down hazard perception into three components: hazard detection, hazard judgment (in terms of interception trajectories) and hazard classification, described as identification in this paper. A number of tests examining only detection and/or judgment have previously been studied and been found to differentiate between novice and experienced drivers in developed countries; for instance, the Deceleration Detection Flicker Test (Crundall, 2009) which has already been applied in Malaysia with mixed results (Lee, Sheppard, & Crundall, 2011), the "What Happens Next?" test (Jackson et al., 2009), and various static hazard perception tasks (Huestegge, Skottke, Anders, Müsseler, & Debus, 2010; Scialfa et al., 2012). Although these tests will need to be validated for use in other countries, they provide a clear direction for future research in Malaysia and other developing countries.



References

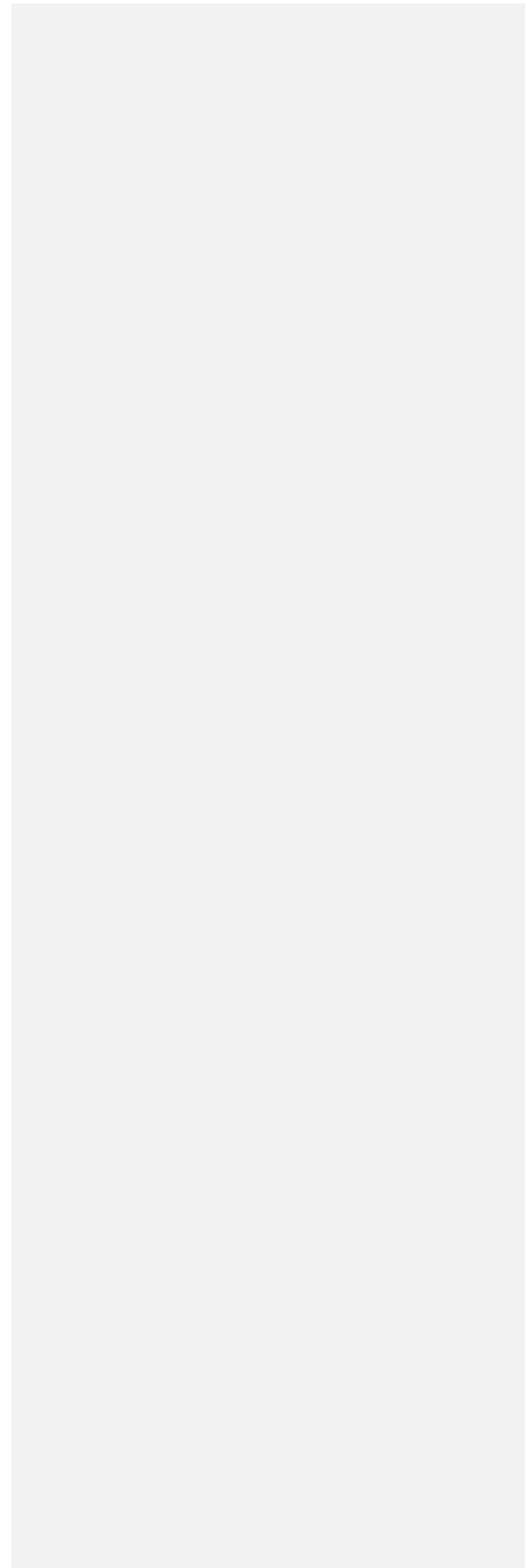


Figure captions

Figure 1. Hazard perception response rate based on verbal identification of pre-defined hazards and a button response made during the hazard window. Panel (a) displays response rates across clip country and whether the hazards were matched across country or not; while panel (b) displays response rates across country clip and driver origin (standard error bars included). Malaysia is abbreviated to MY.

Figure 2. Hazard perception response rate based on verbal identification only (standard error bars included). Malaysia is abbreviated to MY.

Figure 3. Number of extra-hazard responses made for every pre-defined hazard response, depicted across driver origin, clip country and whether the clips were matched or not (standard error bars included). Malaysia is abbreviated to MY.

Figure 4. Number of extra-hazard responses made for every pre-defined hazard response, depicted across driver origin, driver experience and whether the clips were matched or not (standard error bars included). Malaysia is abbreviated to MY.

Figure 5. Average fixation duration. Panel (a) displays average fixation duration across driver origin and clip country; while panel (b) displays average fixation duration immediately prior to and after hazard onset (pre-onset and post-onset), across driver experience (standard error bars included). Malaysia is abbreviated to MY.

Figure 6. Average fixation duration immediately prior to and after hazard onset (pre-onset and post-onset), shown across clip country and whether the hazards were matched across country or not. Standard error bars included. Malaysia is abbreviated to MY.

