

What's the Risk?

A Comparison of Actual and Perceived Driving Risk

Samuel G. Charlton, Nicola J. Starkey, John A. Perrone, & Robert B. Isler

Traffic and Road Safety Research Group, University of Waikato, New Zealand

Abstract

It has long been presumed that drivers' perceptions of risk play an important role in guiding on-road behaviour. The answer to how accurately drivers perceive the momentary risk of a driving situation, however, is unknown. This research compared drivers' perceptions of the momentary risk for a range of roads to the objective risk associated with those roads. Videos of rural roads, filmed from the drivers' perspective, were presented to 69 participants seated in a driving simulator while they indicated the momentary levels of risk they were experiencing by moving a risk meter mounted on the steering wheel. Estimates of the objective levels of risk for the roads were calculated using road protection scores from the KiwiRAP database (part of the International Road Assessment Programme). Subsequently, the participants also provided risk estimates for still photos taken from the videos. Another group of 10 participants viewed the videos and photos while their eye movements and fixations were recorded. In a third experiment, 14 participants drove a subset of the roads in a car while providing risk ratings at selected points of interest. Results showed a high degree of consistency across the different methods. Certain road situations were rated as being riskier than the objective risk, and perhaps more importantly, the risk of other situations was significantly under-rated. Horizontal curves and narrow lanes were associated with over-rated risk estimates, while intersections and roadside hazards such as narrow road shoulders, power poles and ditches were significantly under-rated. Analysis of eye movements indicated that drivers did not fixate these features and that the spread of fixations, pupil size and eye blinks were significantly correlated with the risk ratings. An analysis of the road design elements at 77 locations in the video revealed five road characteristics that predicted nearly 80% of the variance in drivers' risk perceptions; horizontal curvature, lane and shoulder width, gradient, and the presence of median barriers.

Keywords: Risk perception; road design; driving simulator; eye-tracking.

1. Introduction

Since the earliest days of research into driver behaviour, it has been reported that that drivers modify their behaviour according to the risk they perceive (Fuller, 2005; Gibson & Crooks, 1938; Näätänen & Summala, 1974; Taylor, 1964; Watts & Quimby, 1980; Wilde, 1982). It has even been proposed that this behavioural factor is the most important of the three main factors associated with road crashes: behavioural, vehicular, and environmental (Armsby, Boyle & Wright, 1989). Unfortunately, drivers do not always accurately perceive hazards and risks, and as a result, their behaviour may not be appropriate to the circumstances.

In an early study of drivers' perceptions of risk, Pelz and Krupat (1974) showed 60 undergraduate men a 5 minute film of highway driving as seen from the driver's seat and recorded moment-to-moment judgments of risk by means of an "apprehension meter". While watching the film, the participants moved a lever with a scale marked SAFE at one end and UNSAFE at the other according to how safe or unsafe they felt as a driver throughout the film. The participants were divided into three groups based on their driving records: Safe Record, Accidents only, and Violations-or-both groups. The Safe Record group had the highest

baseline level of caution between hazards and the longest duration of elevated caution for each hazard; i.e., the participants in this group recognized driving risk sooner and longer.

Using pairs of still photos taken moments apart, Benda and Hoyos (1983) asked participants to sort 39 different traffic situations (that showed various road and weather conditions) according to their "hazardousness". Their results showed that experienced drivers were able to construct a ranked order of the hazardousness of driving scenes but less experienced drivers (roughly half the years' of experience) grouped the photos according to the type of hazard and did not differentiate the different level of hazardousness shown. The authors also reported that when similar scenes were shown in motion via film clips an equivalent pattern of results was produced and that the participants in each of the experimental conditions tended to separate "comfortable driving" from all other situations. According to Benda and Hoyos, comfortable driving "means driving under good conditions in which drivers do not need to process too much information... relatively few control activities are required. This kind of driving is obviously regarded as fairly nonhazardous" (p. 8). Based on this finding they suggested that drivers' perceptions of hazardousness, or subjective risk, depends on both their amount of experience with various sorts of driving hazards and the information load in the situation, higher information loads leading to higher levels of subjective risk.

Watts and Quimby (1980) asked 60 drivers to make assessments of risk along a 16 mile (25.75 km) route on a rural road and compared the participants' risk ratings to objective risk (calculated from crash data and the participants' speeds). The correlation between the objective risk and the participants' subjective risk was only moderate (Spearman's $\rho = 0.37$) and there were many locations where the risks were underestimated or overestimated. Watts and Quimby suggested that the low levels of perceived risk at some sites may have contributed to the high levels of objective risk, and conversely, there was no crash history at the five locations receiving the highest risk ratings, perhaps as a result of the high levels of perceived risk.

Similarly, Kanellaidis and Dimitropoulos (1994) compared drivers' ratings of subjective risk to the objective risk for five curves on a four-lane divided arterial road in Athens. Thirty-four volunteer drivers drove the 3 km section of road in each direction and subjective risk ratings were given verbally at the midpoint of each curve. Measurement of objective risk at the curves was calculated by filming the curves from both directions and rating the road elements according to the German Guide for Traffic Evaluation of Highways. A very good correspondence between the objective and subjective risk values was observed for the curves (Spearman's $\rho = 0.78$). Two of the curves on the route were generally regarded as accident "black spots" (although no crash history data were provided in the report) and the greatest discrepancy between objective risk and subjective risk ratings occurred at one of these curves.

In a follow-on study, Kanellaidis, Zervas, and Karagioules (2000) used a similar procedure for three different road sections and compared a group of 96 drivers aged 18 to 64 to a group of 40 drivers aged 65 to 75 years. The analysis revealed, once again, that differences between actual risk and perceived risk were associated with increased accident frequency, and that in these cases (where subjective risk is viewed lower than the objective risk) the presence of warning signs becomes most important in maintaining adequate safety margins. The researchers also reported that subjective risk ratings increased with drivers' age, the drivers' familiarity with the roads, and self-assessment of driving skill (the higher the rating of driving skill, the lower the rating of subjective risk).

Groeger and Chapman (1996) showed films of 24 road situations to 64 participants seated at the steering wheel of a partial car. At the conclusion of each scene the participants answered several questions about the level of risk depicted in the scene, how much driving skill was required by a driver in that situation, the amount of control over the danger they would have as

a driver in that situation. Detailed analysis of the participants' ratings indicated that drivers responded to three main characteristics of the situation when considering the road scenes: danger, difficulty, and controllability/abnormality (their level of control and what they would normally expect in that situation). Based on this, and other research, Groeger and Chapman pointed out that although drivers attended to these three factors in their judgements of driving situations, it does not necessarily mean that their judgements of subjective risk were accurate. They argued that, in fact, there is compelling evidence that ratings of subjective risk are highly unreliable and prone to distortions associated with the context in which the judgements are made.

These investigations of drivers' perceptions of risk have shown that drivers do form judgements about the risk of the road and traffic situations they encounter. When compared to the objective risk of the situation, however, the accuracy of those judgements appears to be somewhat variable. Aside from the theoretical interest in the correspondence between drivers' perceptions of risk and the objective risk of various driving situations, there are clear practical reasons for investigating this relationship. Specifically, the published literature suggests that sections of road where drivers' perceived risk is significantly lower than the objective risk (known as risk discordance) may present a significant hazard to drivers (Kanellaidis & Dimitropoulos, 1994; Watts & Quimby, 1980).

There have been very few studies, however, comparing drivers' risk perceptions with some independent estimate of the objective risk, and none of these studies have compared perceived risk and independently estimated objective risk across more than a few driving situations. The goal of the present research was to assess the degree of correspondence between drivers' perceived risk and the objective risk associated with a wide range of road types. Further, the present research used a combination of video, still photographs, and on-road driving at the same road locations and employing the same risk rating scale to explicitly compare differences in methods in stimulus presentation and provide a better understanding of how drivers' risk perceptions are formed. Finally, the collection of risk perceptions across a wide range of road types will allow an analysis of which road features drivers use to judge driving risk and identify hazardous road situations that are under-recognised by drivers (i.e., show the greatest discordance between objective and subjective risk).

2. Experiment 1

2.1. Method

2.1.1. Participants

Seventy-four participants with full (unrestricted) NZ driving licences were recruited to take part in the study, 69 of whom completed the full test protocol (five participants did not complete the testing due to equipment difficulties, eyestrain, or other reasons). The 69 participants (31 males; 38 females) had an average age of 33.7 years (range 17-69 years), and reported an average of 15.7 years since receiving their full NZ licence (range 1.6-50 years). 47.8% of the participants reported that they had never been involved in a motor vehicle crash and 52.2% reported being involved in one or more crash. 75% reported that they had no driving infringements in the past year; 21.7% reported 1 or more. The self-reported ethnicity of the sample was 71% European, 10% Maori, and 18% Other. Ethical approval for the recruitment and test protocols was received from the School of Psychology Research Ethics Committee at the University of Waikato. Each of the participants received a \$20 gift voucher for participating.

2.1.2. Apparatus

Participants were seated in the TARS driving simulator consisting of a complete automobile (BMW 314i) positioned in front of a projection surface on which the test stimuli were projected 2.64 m wide by 2.10 m high (at a resolution of 1920 by 1200 pixels). Details of the TARS simulator have been described elsewhere (Charlton & Starkey, 2011).

2.1.3. Stimulus selection

A set of rural roads from the state highway system were selected such that they were representative of a range of objective risk. The measure of objective risk used was the Road Protection Score from KiwiRAP (the New Zealand Road Assessment Programme). Road Protection Scores (RPS) are calculated for each 100m section of state highway based on a formula incorporating the road design features most likely to influence the occurrence of three primary crash types (run-off road, head-on and intersection) (Waibl, Tate, & Brodie, 2012). Thirteen road design features are rated every 100m and used in the calculation of RPS: lane width, shoulder width, terrain (gradient), horizontal alignment, run-off road risk score, head-on risk score, intersection risk score, left roadside risk score, right roadside risk score, roadside hazard offset left, roadside hazard offset right, roadside hazard severity left, and roadside hazard severity right.

A total of 36 sites were identified on state highways for which RPS scores were available. Each of these sites was selected based on the presence of a particular road characteristic such as lane width, shoulder width, horizontal and vertical curves, and roadside objects such as guard rails, light poles, or drainage ditches. High-definition video (HD resolution, 60 Hz frame rate) of the roads containing the 36 sites were collected from a video capture vehicle driven at the posted speed limits (predominantly 100 km/h) in a safe (i.e., non-aggressive) driving style by an experienced driver. The camera was placed to approximate a driver's perspective while still ensuring the vehicle could be operated in a safe manner. The video collection was repeated several times at each location to ensure that the visibility was clear, weather conditions were comparable, and the presence of other vehicles was minimised (so that the level of risk depicted in the videos arose primarily from the road features rather than momentary conditions or the behaviour of other road users). Each of the videos, with accompanying car and road sounds, was edited into a series of 45 sec "clips" or test stimuli and joined into a 1,690 sec (28 min 10 sec) test video in which each clip was separated from adjacent clips by a 2 sec interval which dissolved from the clip to black and then to the next scene. The resulting video contained a total of approximately 45 km of driving across the 36 locations. Three versions of the testing video were created containing the 36 video clips in one of three random orders, each of them beginning with the same clip (used as a control or "warm-up" to start the video). A 188 sec (3 min 8 sec) training video containing four clips not used in the test stimuli video was created and used during participant familiarisation.

Four clips were cut so that they overlapped in two of the sections of road depicted (two clips had 35 sec of overlap, and two other clips had 15 sec of overlap). These clips were used as a reliability/consistency check on the participants' ratings and they were placed in the videos so that they were never adjacent to one another.

In addition to the test videos, a set of still photos (1920 by 1200 pixels) were created for use as recognition test/risk rating stimuli. Twelve still photos were taken from frames of the test videos, and 14 other photos taken from the raw video (not contained in the 36 test clips or at locations near them). Two of the "new" photos were designated as practice stimuli, and the remaining 24 photos were put into three random orders for presentation to participants. One

of the photos taken from the test video was used as a reliability check and was repeated three times in each presentation order (in nonadjacent locations).

2.1.4. Procedure

Following informed consent and completion of a brief demographic questionnaire the participants were seated in the simulator and shown the 3 min training video. During the training video the participants were asked to engage in a steering task in which they were instructed to use the steering wheel to keep a yellow dot lined up on the centre of their lane front of them (the same secondary steering task as used by Charlton, 2006). The steering task was introduced to make the visual and attentional demands of the situation more representative of driving a car. During the test video the yellow dot was removed from the screen so as not to produce a distraction although participants were still instructed to “steer” appropriate to the road direction. During both the training and test videos participants were also instructed to provide moment-to-moment judgments of driving risk by means of a risk meter (an analogue to Pelz & Krupat’s apprehension meter, 1974). The risk meter consisted of a thumbwheel mounted on the right side of the steering wheel. The thumbwheel controlled a pointer which moved along an on-screen scale anchored with the words “Safe” at the bottom and “Unsafe” at the top with nine calibration lines between (see Figure 1) and recorded the participants’ momentary risk ratings as a continuous variable. The instructions to the participants were to report how safe or unsafe they felt as the driver with the “Safe” end of the scale referring to feeling completely at ease such as while being at rest or parked while “Unsafe” referred to feeling extremely threatened or in immediate danger of being involved in a serious accident. Participants were also asked to move the risk meter all the way down to the “Safe” end during the 2 sec dark intervals separating the different scenes.



Figure 1. A scene from the test video showing the on-screen risk scale.

After completion of the training video, the participants were shown the full 28 min test video (one of the three orders) during which they provided moment-to-moment risk ratings for the 36 video clips and “steered” the car in the appropriate direction (albeit without the yellow dot target). At the end of the video the participants were invited to take a short (2-3 min) rest break and get out of the car to stretch. The participants were then given instructions about the self-paced recognition and risk rating task for the still photos. Participants were asked to indicate whether they recalled seeing each photo location in the video they just watched using the turn indicator (moving it up for yes, down for no) and then rating the risk of each location using the risk meter. The series of 28 photos (two practice photos, 24 test photos, and two reliability check photos) began with an introduction to the two practice photos and the presentation rate

was self-paced (participants advanced to the next photo by moving the headlight stalk). After completing all of the ratings the testing was concluded and the participants were thanked for their participation and given a \$20 gift voucher.

2.1.5. Data analysis

For analysis of the video ratings, 35 specific points were identified beforehand (1 location per clip, excluding the warm-up clip) based on the presence of particular road features of interest (e.g., poles, ditches, vertical or horizontal curves). Another 22 points were identified from the participants' risk ratings (where ratings were high or large changes in ratings occurred). Points of interest that included other road users such as vehicles parked at the roadside, oncoming trucks, and pedestrians were excluded. In addition there were 20 locations depicted in the videos that were selected on the basis of their high RPS (objective risk), yielding a total of 77 locations of interest/points in the video for the analysis of participants' risk ratings. The locations depicted in the 12 still photos were all drawn from the initial 35 points of interest in the video clips such that the road types and situations were representative of the range of locations shown in the video clips.

The risk ratings were analysed in three ways: comparison of the ratings of perceived risk with the objective risk, comparison of the ratings of the video with the still photos, and assessment of the consistency of the participants' ratings by comparing the repeated video clips and photos.

2.2. Results

Participants' risk ratings during the videos ranged the entire risk rating scale (where "Safe" = 1 and "Unsafe" = 10); the point in the video with the highest risk rating had a mean of 6.68 (SD = 2.60) and the point with the lowest risk rating had a mean of 1.93 (SD = 0.91). Risk ratings for the 24 photos showed a very similar pattern, with the mean of the participants' ratings ranging from a high of 5.96 (SD = 2.19) to a low of 2.06 (SD = 1.18). No significant differences in risk ratings were identified as resulting from participant gender or age.

2.2.1. Comparison of perceived and objective risk

The relationship between the participants' ratings of risk and the objective risk for the 77 locations of interest in the video can be seen in Figure 2. In the left panel of the figure are the mean risk ratings plotted against their corresponding RPS scores for all 77 locations of interest. The correlation between risk ratings and RPS scores overall was rather low; Pearson's $r(68) = .081, p = .486$. Because of the way RPS scores are calculated from objective risk components, the 10 points with high RPS scores were all located at intersections. In the right panel of the figure are the mean risk ratings and RPS scores for the 57 locations of interest excluding the 20 intersections. The correlation for this subset of locations was strong and positive; Pearson's $r(68) = .771, p < .001$. In other words, the participants' risk ratings agreed quite well with the objective risk scores, with the exception of the ratings of intersections, which although having high risk RPS scores, were not rated as high risk by the participants.

The right panel of the figure also shows some locations where the risk ratings were above the 95% confidence intervals, and others that fell well below the confidence intervals. The locations lying above the interval represent situations that the participants perceived as riskier than the objective risk in the situation. All of the points associated with over-rated risks were instances of narrow lane width, horizontal curves, and wire rope barriers. In contrast, every location with narrow road shoulders, ditches, and poles in close proximity to the road were under-rated as risks by the participants. The remaining four (of 17 total) under-rated risks included banks on the roadside, wide (dual) centre lines, and one location with a grass median (but no barrier).

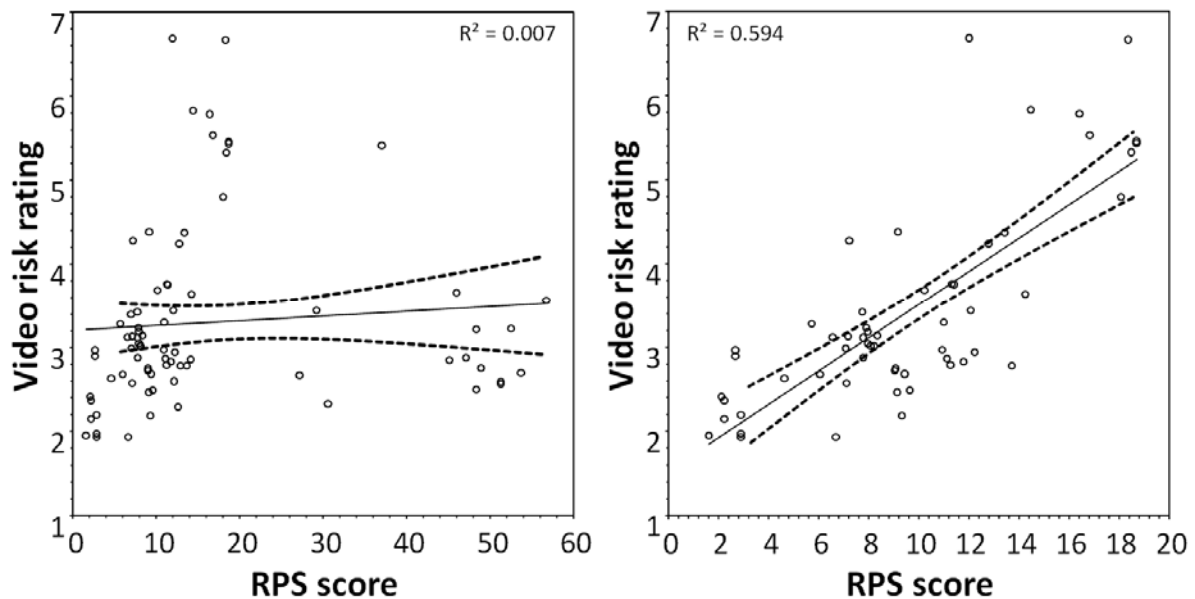


Figure 2. Mean risk ratings and RPS scores for video scenes including (left panel) and excluding the 20 intersections (right panel). Dashed lines show 95% confidence intervals.

A multiple regression analysis predicting the mean risk ratings for each of the 77 locations from the 13 components of the RPS scores at those locations revealed that the single largest predictor of the participants' ratings was the horizontal alignment score; $\text{Adj } R^2 = .532$, $F(1,73) = 85.00$, $p < .001$. The best combination of RPS components predicting the risk ratings was horizontal alignment, lane width, shoulder width, terrain (gradient), and right roadside risk; $\text{Adj } R^2 = .785$, $F(1,69) = 54.97$, $p < .001$. These 5 RPS components together accounted for nearly 80% of the variance in the participants' risk ratings; none of the other component measures were significant predictors of the risk ratings. (It should be noted that in this context the RPS component right roadside risk was essentially a measure of whether or not the opposing lanes were separated by a physical barrier.)

2.2.2. Comparison of video and still photo risk ratings

Figure 3 shows the risk ratings from the 12 photos presented to participants in the second stage of the driving simulator testing compared to the risk ratings for the corresponding locations in the video. As can be seen in the left panel of the figure, the correspondence is very high, statistical analysis indicated a strong positive correlation; Pearson's $r(68) = .926$, $p < .001$. Two locations fell above the 95% confidence interval indicating that they appeared riskier in the photos than they did during the video and, as shown on the right of the figure, both locations were bridges.

Overall, there was a slight tendency for the participants to rate the 12 still photos as being riskier than the corresponding locations in the video ($M = 3.74$, $SD = 1.18$; $M = 3.52$, $SD = 1.45$; for photos and video respectively). A repeated measures analysis of variance indicated that this difference in media type was statistically reliable; $F(1, 67) = 9.48$, $p = .003$, $\eta_p^2 = .124$. There was also, however, a statistically significant interaction between media type and the individual locations, indicating that the correspondence between the ratings given to the stills compared to the video differed somewhat from location to location; $F(11, 737) = 11.37$, $p < .001$, $\eta_p^2 = .145$. Analysis of the individual locations revealed that this was restricted to three of the 12 locations; a location with narrow lanes and a wooden edge rail (that was rated riskier in the video), and a bridge (shown as location 7 in Figure 4) and a vertical crest (both rated riskier in the photos than in the video). Comparing the participants' ratings of 10 of the

photos to the RPS scores for those locations (RPS scores were unavailable for two locations) produced a strong positive correlation, Pearson's $r(68) = .847$, $p = .002$; only slightly lower than the correlation between video ratings and RPS scores for the same locations; $r(68) = .886$, $p < .001$.

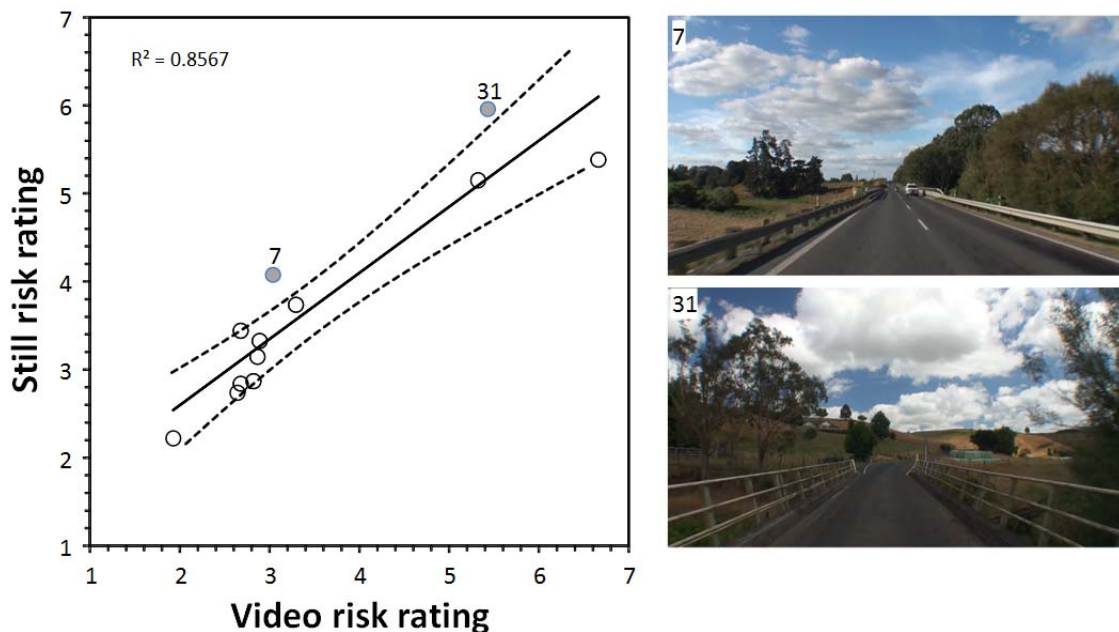


Figure 3. Mean risk ratings for 12 still photos and corresponding video locations of the roads over-rated in the still photos. Dashed lines show 95% confidence intervals.

2.2.3. Consistency of the risk ratings

The consistency of the participants' ratings was examined by comparing the ratings at the two locations that were depicted in overlapping video clips, as well as the ratings obtained for the photo that was presented three times during the second stage of simulator testing. The ratings for the repeated video scenes were very consistent, showing nearly identical reactions to the road environment on successive viewings, even though the locations were presented at different points in the two clips, and the clips themselves came at different points in the videos. The largest difference in any single participant's risk ratings on successive viewings was 2.21 for the two clips with 15 sec of overlap (the difference in ratings averaged across all participants was 0.282). For the two clips with 35 sec of overlap the largest individual difference in ratings was 3.71 (the mean across all participants was 0.497). A consistency check for the ratings of the two clips with 35 sec of overlap calculated across all 69 participants at the midpoint of the overlapping section in produced a highly significant Cronbach's Alpha of .904 ($M = 2.45$, $SD = .145$; $M = 2.54$, $SD = .163$ for the two points). The two clips with 15 sec of overlap also displayed a nearly identical pattern, with a Cronbach's Alpha of .888 calculated at the midpoint of the overlapping section ($M = 2.06$, $SD = .127$; $M = 2.01$, $SD = .118$ for the two points). In the still photos presented during the second stage of the simulator testing, the reliability check photo received very similar ratings each time it was presented ($M = 2.87$, $SD = .147$; $M = 3.16$, $SD = .169$; $M = 2.94$, $SD = .163$ for the first, second, and third viewings). A Cronbach's Alpha indicated high consistency in the ratings across the three viewings ($= .925$).

3. Experiment 2

A second sample of participants was recruited to have their eye movements recorded while rating the risk of the video and still photo stimuli. The purpose of this was to provide information about where participants looked when viewing the videos and photos, and which

features of the driving situations attracted their gaze. Also of interest was the identification of potential implicit measures of risk such as saccades, pupil dilation, and eye blinks by comparing them to participants' explicit risk ratings.

3.1. Method

3.1.1. Participants

A total of 10 participants were recruited to take part in the eye-tracking testing. Ethical approval for the recruitment and test protocols was received from the School of Psychology Research Ethics Committee at the University of Waikato. Each of the participants received a \$20 gift voucher as a token of appreciation. The 10 participants (4 males; 6 females) had an average age of 24.4 years (range 19-30 years), and reported an average of 8.4 years since receiving their full NZ Licence (range 4-13 years). Four of the participants reported that they had no crashes and six reported being involved in one or more crash. Only one of the nine drivers reported that they had received a driving infringement in the past year. The self-reported ethnicity of the sample was 7 European, 1 Maori, and 2 Other.

3.1.2. Apparatus

An SR Research EyeLink II eyetracker set for monocular recording (participant's preferred eye) at a sample rate of 250 Hz was used to monitor eye movements whilst participants viewed a flat-panel display screen (93cm x 52cm, 1920 x 1080 pixels) from a distance of 90cm (field of view = 55° horizontal x 32° vertical). The tracker recorded eye position data, blinks and pupil size. The position data was smoothed with a low-pass filter (-3dB at 32 Hz) and x and y eye velocity (V_x , V_y) was calculated by applying a low-pass differentiator (-3dB at 32 Hz) to the original (x, y) position data. Saccades were detected in the velocity traces using a robust saccade-detection algorithm (Liston, Krukowski & Stone, 2013). The stimulus presentation software and on-screen risk scale were the same as used in the driving simulator testing. The risk scale was superimposed over the right hand side of the video display (15° to the right of the centre of the road) and provided visual feedback to the participants as to their risk rating. Eye position data was analysed for points of interest in the central or left region of the screen only, to avoid the saccades or fixations towards the on-screen risk rating scale.

3.1.3. Procedure

The same informed consent, participant instructions, and training procedure used in the driving simulator testing was used for the eye-tracking testing. As with the driving simulator testing, the participants rated the perceived risk of the road video using a thumbwheel attached to a steering wheel in front of them. The participants were asked to engage in the secondary steering task, with the dot present during training, and the dot absent during the test video. To minimise discomfort (the eye-tracking equipment becomes uncomfortable if worn for extended periods), a subset of 20 of the video clips used in the driving simulator testing was selected for the eye-tracking testing. The clips were selected to provide a range of different road types and hazards. These clips, plus the same initial clip used for practice, resulted in a test video 984 sec (16 min 24 sec) in length. The same 3 min 8 sec training video containing 4 clips used in simulator testing was used for instructing the eye-tracking participants. The same set of 24 photos (and 2 practice photos) used for the driving simulator testing were used for the eye-tracking testing. When the testing was concluded the participants were thanked and given a \$20 gift voucher.

3.1.4. Data analysis

The eye-tracker data were analysed using custom software (MatLab R2013b, Mathworks) and the blink and saccade totals for each participant and movie clip were extracted. The blink lengths (msec) were also calculated from the blink onset and offset times.

Five analyses of the eye movement data were conducted: comparison of the eye tracking risk ratings to the risk ratings from the driving simulator testing; eye-tracking and fixations for under-rated risks; number and duration of blinks in relation to risk ratings; comparison of pupil size and saccades to risk ratings; and the spread of participants' eye fixations when looking at the still images.

3.2. Results

3.2.1. Comparison of risk ratings

To confirm that the participants in the eye tracking experiment were perceiving the same level of risk as those in the driving simulator testing, we compared the average rating for each of the 20 locations in the test video against the average rating for the corresponding locations in the Experiment 1 driving simulator testing. (The risk rating data from one of the participants was not recorded because of a computer software problem and so only 9 participants' data were included in this comparison.) As shown in the left panel of Figure 4, the correspondence between the risk ratings obtained for the two types of testing was high and statistical analysis indicated a strong positive correlation; Pearson's $r(19) = .778$, $p < .001$ for the video ratings. The overall average of the risk ratings for the two test protocols were very close ($M = 3.41$ for the driving simulator, $M = 3.49$ for the eye-tracking test), although the variability was higher in the simulator ($SD = 1.15$, $SD = 0.76$, for the simulator and eye-tracking test respectively), possibly due to the relatively small sample of participants in the eye tracking experiment, or possibly the somewhat different age and ethnicity make-up of that sample.

The right panel of Figure 4 shows the comparison of the mean risk ratings for the 24 still images used during the two experiments. Once again, the correspondence between the risk ratings obtained for the two types of testing was high and statistical analysis indicated a strong positive correlation; Pearson's $r(23) = .947$, $p < .001$ for the risk ratings of the still photos. The overall risk ratings of the photos in the two test protocols were nearly the same ($M = 3.72$ for the driving simulator, $M = 3.92$ for the eye-tracking test), although the variability was higher in the eye-tracking experiment ($SD = 1.01$, $SD = 1.19$, for the simulator and eye-tracking test respectively).

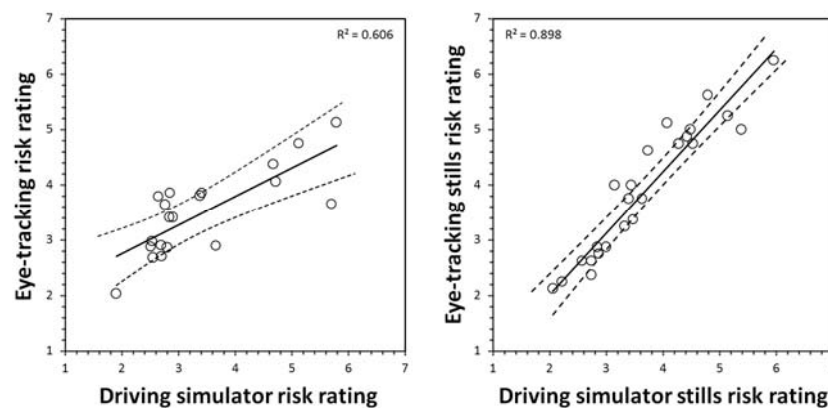


Figure 4. Mean risk ratings for 20 locations in the eye-tracking video and corresponding locations in the driving simulator experiment video (left panel) and the 28 still photos in each experiment (right panel). Dashed lines show 95% confidence intervals.

3.2.2. Eye-tracking and fixations for under-rated risks

The results of the driving simulator experiment showed that the participants rated roadside hazards such as power poles and ditches as low risk even though the objective risk (RPS) is high. A useful question to ask is whether or not drivers noticed these hazards or attended to them. To answer this question, the participants' eye-tracking and fixations were analysed to determine if the participants looked in the direction of the poles. If they did not look at them, then it may suggest that they did not think them sufficiently important to fixate or to pursue with their eyes.

To address this question, a video clip containing a straight piece of road with a series of power poles close to the edge of the road was selected for detailed analysis. In the video clips, these move from the centre of the screen to the left edge over a period of about 2.5 sec as the driver approaches and passes a pole (see Figure 5). For each pole the X location at the start and end of a series of 2.5 sec segments from the video was located.



Figure 5. Screen positions of a roadside pole used in the analysis of participants' eye-tracking and fixations for under-rated risks.

An example of pole locations for three of these video segments is shown as bold diagonal lines in the three panels of Figure 6. For these three example segments, the position of the poles were measured as beginning at image pixel location 866 and ended up at a point 32 pixels from the left edge of the screen. (For purposes of this analysis, all positions to the right of the centre of the screen were set to 0 in order to include only those eye movements on the same side of the road as the pole). The mean X eye location across all 10 participants over these 2.5 sec segments of the video clip were calculated from a series of 4 msec time samples, along with the 95% confidence intervals for the mean as shown in Figure 6. As can be seen in the figure, except for the very start of the segment, where the pole was close to the centre of the road in the movie image, on average the participants' gaze did not include the position of the pole.

Roadside ditches were also underrated (relative to RPS scores) by the participants. We carried out a similar analysis to that used for the power poles and examined participants' average fixation locations relative to the position of a ditch to determine whether or not they fixated roadside ditches. The location of a ditch in the video requires both an X and Y position to define and it is not possible to simply compare the eye fixations relative to a line. For this analysis the average (x, y) location across a 2.5 sec segment of a video clip containing a prominent ditch at the side of a straight section of road was calculated.

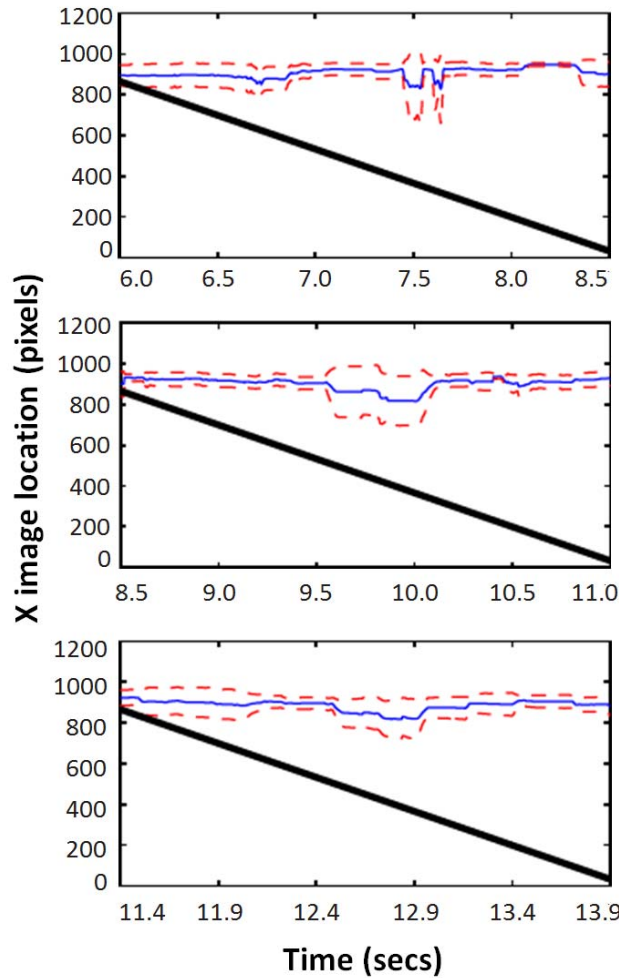


Figure 6. Participants' mean eye fixations (solid line) compared to roadside pole locations (bold diagonal line) for three video segments. Dashed lines indicate 95% confidence intervals for participants' eye fixations.

Figure 7 shows the 10 participants' mean eye fixations superimposed on an image from the middle of the video segment distribution relative to the ditch location (indicated with the dashed oval in the figure). As can be seen in the figure, none of the participants made eye fixations in the location of the roadside ditch.

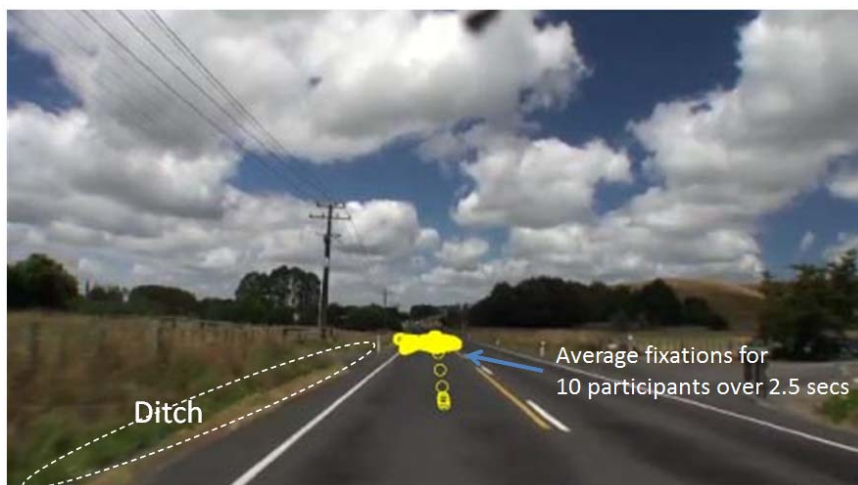


Figure 7. Participants' mean eye fixations (circles) compared to a roadside ditch location (dashed lines) for a 2.5 sec video segment.

To investigate whether the above results could be a consequence of the participants simply not fixating any roadside objects, a similar 2.5 sec section of straight road containing a roadside object other than a pole or ditch was selected from the video. The roadside object was an advance curve warning with a supplementary speed advisory (see Figure 8). Over the course of the 2.5 sec video the sign moved from location 853 to 613 pixels from the left of the screen. The mean X eye position for the participants relative to the position of the roadside sign across the 2.5 sec video is shown in the figure, and as can be seen, a substantial number of eye fixations did occur to the roadside area containing the sign. (In this example the car was travelling slower and so the bold diagonal line indicating the sign position is less steep than in the roadside pole examples). From this example it is apparent that the participants' functional field of view did include roadside objects, however, the roadside poles and ditch did not attract fixations to the degree that the example curve advisory sign did.

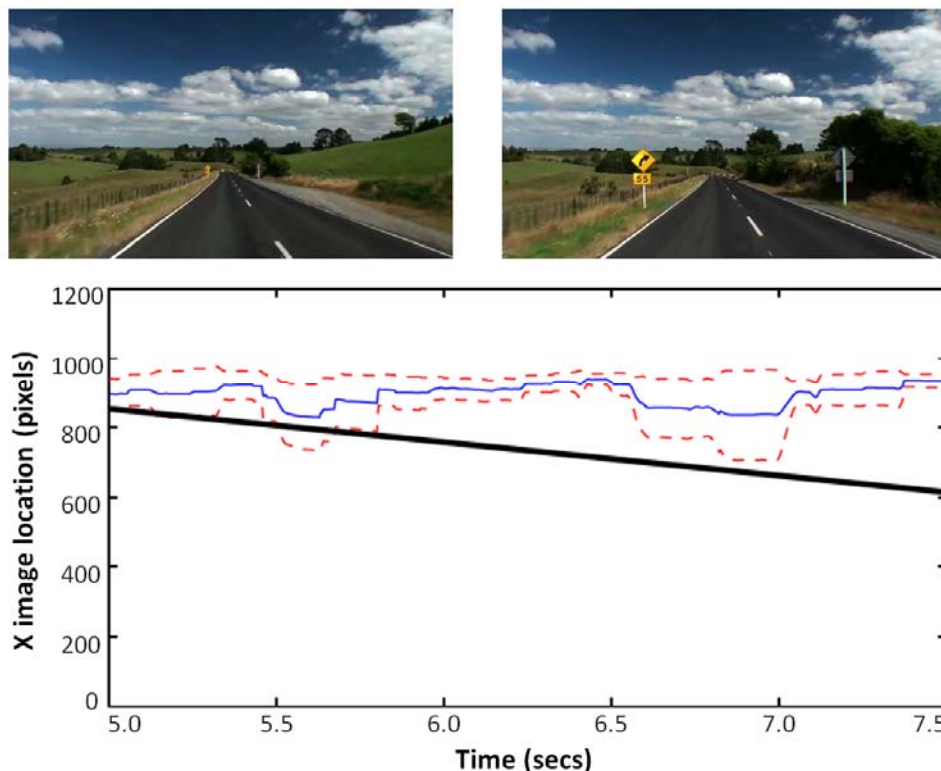


Figure 8. Participants' mean eye fixations (solid line) compared to roadside sign locations (bold diagonal line).

Dashed lines indicate 95% confidence intervals for participants' eye fixations.

3.2.3. Analysis of blinks

Given previous reports that blink rate may be a good indicator of attentional demands (cognitive workload) and stress (Brookhuis & de Waard, 2010), it was of interest to explore the relationship between blinks and subjective risk in the current study. First, each participant's total number of blinks was counted and the duration of each blink was summed for each of the 20 clips in the video. Linear regression analyses were then conducted for each of these variables against the average risk rating for the clip (over the total length) to see if there was any relationship between the number of blinks and the total duration of blinks (total eye closure time) and the risk ratings. Statistical analysis indicated that both the number and total duration of the blinks decreased significantly as the risk rating increased; Pearson's $r(19) = -.534$, $p = .015$ for number of blinks and $r(19) = -.553$ for total blink duration ($p = .011$).

In order to examine whether or not blinks corresponded to perceptions of risk at specific locations, as opposed to averaging across the 45 sec duration of each video clip, each video clip was divided into 'Low risk' (where the risk rating was ≤ 5) and 'High risk' (> 5) segments. For each segment the total blink length (i.e., total eye closure time) was used to calculate the blink density (total eye closure time/ length of segment) for each of the high and low risk locations in the video. This measure gives an indication of the total blink duration that occurred during the period when the participant rated the video clips as low or high risk normalised for the length of time they considered the clip segments to be low or high. (One participant had no rating data and another never rated any segment of the clips as being 'High', leaving 8 participants for this analysis). A repeated measures t-test comparing blink density during the High and Low risk sections of the clips was not significant; $t(7) = 1.7, p = .13, d = 0.600$; but the effect size (d) suggested a possible relationship existed (and that the failure to meet the critical t value may have resulted from the small sample size).

3.2.4. Analysis of pupil size and saccades

For each of the 20 video clips, the mean pupil size ($n = 10$ participants) and mean number of saccades during each video clip were calculated and, similar to the eye blink analysis, these measures were compared to the average risk rating over the total length each clip ($n = 9$ participants). The results of this analysis indicated that the participants' pupil size and number of saccades were both positively correlated with ratings of risk; Pearson's $r(19) = .531, p = .016$ for pupil size and $r(19) = .737, p < .001$ for number of saccades.

3.2.5. Analysis of eye fixations for still images

For each of the 12 still images containing target scenes (present in the video clips), the location and duration of participants' eye fixations were represented as a 'heatmap'; the (x, y) fixation data were smoothed using a 2D Gaussian ($SD = 20$ pixels) with an amplitude equal to the fixation time (in msec). These heatmaps give an indication of how long the participants spent looking at each part of the image. (It should be noted that this was done for each participants' fixations excluding values corresponding to the location of the on-screen risk meter scale.) Figure 9 shows the resulting heatmaps produced for a still image rated as low risk (top panel) and an image rated high risk. As can be seen in the figure, the spread of fixations appeared to be broader for the images rated as high risk, particularly the horizontal spread. In order to assess this relationship statistically we fit a 3D Gaussian to the average heatmap distribution from all participants and recorded the SD of the fitted Gaussian. The horizontal spread (xSD) extracted from each image was correlated with the mean risk ratings for those images. The results of this analysis suggested that the horizontal spread of fixations was indeed positively correlated with ratings of risk; Pearson's $r(11) = .571, p = .053$.

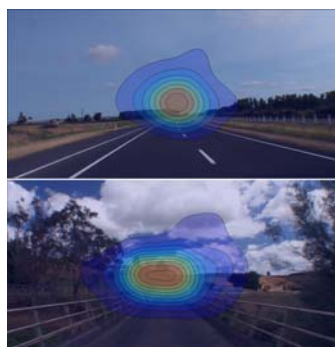


Figure 9. Gaussian 'heatmap' indicating the average of all participants' fixation patterns for the still images. The top panel shows a still image rated as low risk, the lower panel shows an image rated as high risk.

4. Experiment 3

Another sample of participants was recruited to drive an actual car over a subset of the roads contained in the videos presented to participants in Experiments 1 and 2 and provide verbal risk ratings at 13 selected points of interest. This was followed by an interview session in which participants watched a video recording of their drive and provided reasons for their risk ratings. The risk ratings, and the participants' reasons for their risk ratings provided another point of comparison and cross-validation, as well as an explicit narrative to aid in interpretation of the data collected.

4.1. Method

4.1.1. Participants

A sample of 15 participants over 25 years of age, who held a full New Zealand driving licence was recruited to take part in the eye-tracking testing. Participants were recruited via flyers posted on noticeboards around the University, local community centres, shops and cafes and social networking sites. Ethical approval for the recruitment and test protocols was received from the School of Psychology Research Ethics Committee at the University of Waikato. Fourteen participants (8 males, 6 females, average age = 41.6 years, range 25-50 years, 10 of European descent and 4 'Other') completed all or part of the drive (the remaining participant was recruited to trial the procedure). Of the 14 participants who took part in the study, data was incomplete for three participants due to poor weather, other time commitments and equipment failure. The participants had been licensed drivers for 23.3 years on average (range 2-33 years). Six drivers reported that they had never been involved in a crash; eight had been involved in at least one crash. Three drivers had received an infringement notice in the past year. Each of the participants received a \$40 gift voucher at the end of the drive and another \$10 voucher at the end of the interview session.

4.1.2. Apparatus

A Suzuki SX4 vehicle, 2012 was fitted with two video cameras (HD quality, equipped with a standard lens, $f=55$ mm); the first was attached via suction cups to the front windscreen, recording the road scene ahead; the second was attached to the small window on the passenger side of the vehicle to record the driver's behaviour. A laptop computer generated a tone to prompt the participant to provide a verbal risk rating at predetermined GPS coordinates (points of interest) along the route. The computer also synchronised the video files from the two cameras and stored them.

During the post-drive interview participants viewed edited videos of their drive on a flat-panel display screen (93cm x 52cm, 1920 x 1080 pixels) from a distance of 2.3m. For each participant, a 35 sec sections of the video containing the locations of their risk ratings 25 seconds before the tone to 10 seconds after) were extracted from the full video of their drive.

4.1.3. Procedure

The participants drove a route containing 13 of the points of interest shown in the driving simulator in Experiment 1. The route (approximately 180 km round trip) was selected as it was close to the research base at the University of Waikato, contained roads with a variety of features and hazards, and the drive could be completed within a reasonable period of time (approx. 2 - 2.5 hr). Participants were given a break and provided with a drink at the mid-point of the drive.

After contacting the researchers, each participant was sent an information sheet about the study and invited to meet a researcher at a mutually convenient time. On arrival at the laboratory, the purpose of study was explained, the participant was shown a map of the route, any questions

the participant had were answered, and the participant was then asked to sign a consent form and complete a demographic and driving history questionnaire. Participants were asked to comply with all normal roads rules and regulations and drive as they would in their own car. Each participant was then taken on a short test drive (15-20 mins) to familiarise them with the vehicle and the tone that prompted them to provide a verbal risk rating. The verbal ratings of risk were recorded (using pen and paper) by the research assistant in the vehicle in addition to being captured as part of the video recording of each drive. The risk rating scale and definitions were the same as that used during Experiment 1.

The participants were invited back to the laboratory within 5 days to view the video of their drive. During this session participants were instructed that they would be shown two practice video clips, followed by clips from their drive at the locations where they were prompted to provide risk ratings. Each clip was preceded by a 5 second on-screen countdown and ended with a blank screen. At the end of each clip participants were given the opportunity to change the risk ratings they gave during their drive (or leave them unchanged) and comment on road features that contributed to their ratings. The interview sessions took approximately one hour and were recorded (audio and video) for subsequent analysis. The participants' comments were reviewed and post-hoc categories of road features contained in the videos and/or mentioned in the comments (curves, visibility, traffic, terrain, narrow road, signs, straight road, bridge, road markings, junction, banks, weather, no shoulder) were derived by two scorers. The number of comments in each category were counted, regardless of whether participants mentioned the particular road feature to justify a high or low risk rating (the focus was on identifying features than informed risk ratings generally, rather than focusing on high risk features only).

4.2. Results

4.2.1. On-road risk ratings

The mean risk ratings were lower for the on-road drive ($M = 3.44$, $SD = 0.97$) compared to the ratings obtained for the same roads viewed in the driving simulator ($M = 4.03$, $SD = 1.58$), but this difference was not statistically reliable; $t(24) = -1.159$, $p = .258$. Figure 10 shows a comparison of the participants' mean risk ratings for the 13 road segments compared to the mean risk ratings from the corresponding video clips during Experiment 1. As can be seen in the figure, there was a strong positive relationship between the risk ratings from the on-road drive and the video-based ratings; Pearson's $r(13) = .791$, $p = .001$. Also shown in the figure, most data points were within, (or at) the 95% confidence intervals, except for the ratings of two test locations. For the first of these locations (a narrow road with a white guard rail on the left and a high bank on right), the on-road rating was lower than that predicted from the video-based risk rating. For the second location (straight road with narrow shoulders and roadside power poles), the on-road rating was slightly higher than predicted. The high correlation between the video-based ratings and those obtained from the naturalistic drive support the pattern of the subjective risk data collected in the driving simulator sessions and suggest that these ratings correspond well to perceptions of risk while driving on the road.

4.2.2. Post-drive interview

The mean risk ratings from each point of interest on the naturalistic drive ($M = 3.44$, $SD = .97$) were remarkably similar to those obtained from the same participants during the laboratory video viewing ($M = 3.41$, $SD = .97$). A paired samples t -test confirmed that there was no significant difference between the on-road and video-based ratings, $t(12) = .50$, $p = .63$, suggesting that the level of risk experienced was similar across both presentation modalities.

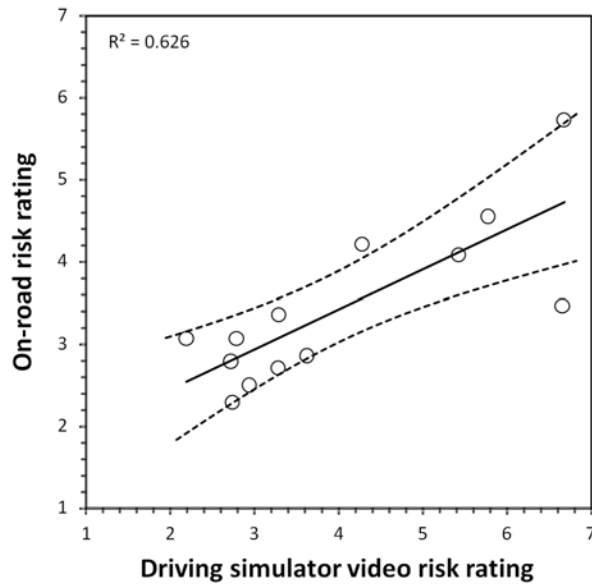


Figure 10. Mean risk ratings of the 13 test roads from the on-road and the video-based driving simulator ratings (1 = ‘safe’, 10 = ‘very unsafe’). Dashed lines show 95% confidence intervals.

In regard to the road features that participants reported as contributing to their risk ratings (Figure 11) the most commonly mentioned related to curves (e.g., *swerving corner*), visibility (e.g., *can't see ahead*), traffic (e.g., *idiot on motorbike; oncoming traffic*) and terrain (e.g., *brow of hill*). The road width was also noted, as were speed advisory signs, particularly heading into curves (e.g., *35 km/h sign close to corner*). All of the participants commented that the presence of the single lane bridge influenced their risk ratings (it was only present in one clip). Road markings (e.g., *yellow lines; no overtaking*), junctions, banks, weather and the lack of a shoulder were mentioned less frequently. The presence of poles and ditches did not appear to contribute to the participants' risk ratings.

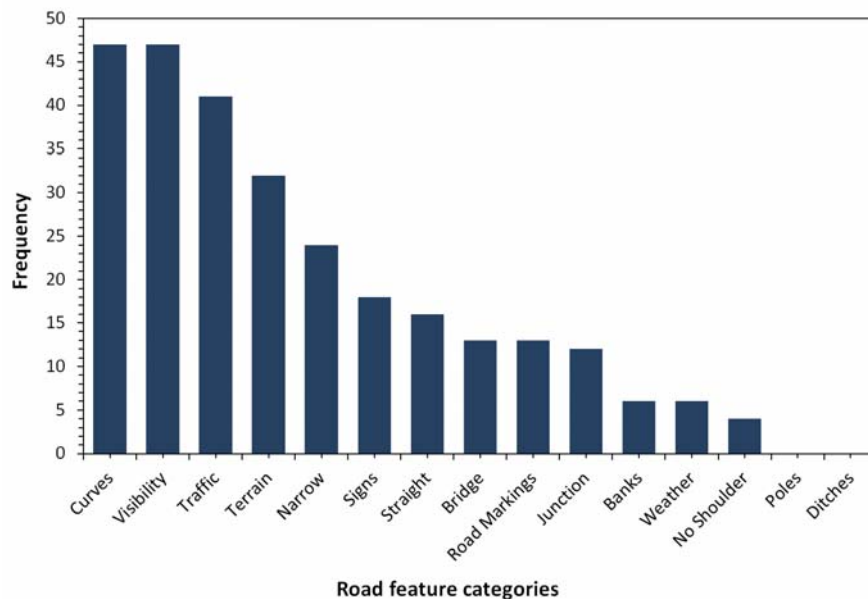


Figure 11. The frequency of reporting of specific road features as contributing to participants' risk ratings.

5. Discussion

The findings of the present study indicate that drivers do experience a wide range of risk levels for different road situations. This finding is in line with previous research indicating that drivers can and do form judgements about driving risk (Fuller, 2005; Groeger & Chapman, 1996; Pelz & Krupat, 1974; Watts & Quimby, 1980). The regression analysis of risk ratings from the driving simulator revealed that the participants used curves, hills, road width, and the presence or absence of a divided median to judge the risk of a road. These features alone predicted nearly 80% of the variation in the participants' ratings and, notable by their absence, none of the other road features making up RPS scores contributed to drivers' perceptions of risk. This finding is similar to the four road engineering factors predicting drivers' ratings of risk reported by Kanellaidis (1995): separation of opposing traffic, cross-section characteristics, alignment, and signing.

In the present study the level of risk experienced was, in general terms, commensurate with the objective risk for the roads. In other words, locations with high objective risk were correctly perceived as high risk and locations with low objective risk were perceived as low risk. As with previous studies, however, there were some notable exceptions; situations where the participants' perceived risk either over- or under-estimated the objective risk. Horizontal curves and narrow lanes were associated with over-rated risk estimates, while intersections and roadside hazards such as narrow road shoulders, power poles and ditches were significantly under-rated. The finding that drivers rated the risk associated with curves as high (and in some cases over-rated the risk) is perhaps not surprising given the previous findings that have shown horizontal alignment to be an important dimension of drivers' perceived risk (Kanellaidis & Dimitropoulos, 1994) and that most, if not all, curves are accompanied by warning signage (Charlton, 2007). Similarly, the finding that drivers view bridges and narrow lanes as high risk has also been reported previously (Watts & Quimby, 1980) and may in part be due to an increase in perceived speed at these locations (Lewis-Evans & Charlton, 2006). In the present study, the presence of wire rope barriers and bridge rails tended to increase participants' ratings of perceived risk, particularly for the still photos in the case of bridge rails. The data collected during the post-drive interviews supported these findings, with curves, terrain and road width being three of the most commonly reported features that influenced drivers' risk ratings.

The finding that drivers under-rated the risk associated with roadside hazards such as power poles and ditches (as compared to their objective risk) has not been previously studied or reported to the same extent. It has been suggested, however, that drivers appear to focus on factors contributing to the risk of crash occurrence rather than risk factors associated with the severity of a crash. For example, Lund and O'Neil (1986) proposed that behavioural adaptation would be more likely to occur for changes that affect crash probability (studded tyres) than changes that affect injury probability (seat belts). They argued that this was because changes that reduce the likelihood of a crash also often provide more direct and immediate feedback to the driver (as compared to injury-reducing measures) (Lund and O'Neil, 1986). In the case of the present study, the road locations containing power poles and ditches can be thought of as severity magnifiers rather than directly contributing to crash occurrence. The RPS scores correctly take into account the additional risk posed by these hazards, but drivers apparently do not view their presence as contributing to driving risk.

The finding that intersections were also significantly under-rated by our participants requires a different explanation. The potential for conflict with other vehicles at intersections does contribute to crash likelihood rather than magnifying the severity of crashes. Intersections were mentioned by some of the participants during the interviews as contributing to their risk ratings, particularly when other vehicles were present, but the risk ratings remained moderately low.

Unlike roadside ditches and poles, which were not mentioned by any of the participants, the risks associated with intersections were under-rated rather than simply not considered. It may be that because intersections do not place any additional demands on drivers' vehicle control skills they do not contribute to an increased sense of risk. Using a similar risk rating procedure, Groeger and Chapman (1996) found that perceived difficulty and degree of controllability were important constituents of drivers' feelings of risk. Likewise, Fuller (2005) has suggested that driving difficulty may be conflated with feelings of risk to some degree. In the present experiments, the risk ratings of the still photos were slightly (but significantly) higher than those from the videos, which in turn were slightly higher than those from the on-road experiment. The different sense of controllability associated with these methods (low controllability for photos and high for on-road) could help account for these differences between the methods.

It must be acknowledged that the use of still photos or videos in a driving simulator to assess risk perception does have potentially significant limitations inasmuch as the sense of the risk of a real-life crash occurring cannot be conveyed fully in these simulated circumstances. The pattern of results described above, however, (lower ratings for on-road procedures, and higher for still photos) would tend to discount the degree of these limitations. The high level of consistency in the risk ratings from the complementary methods used in the present research (video, photo, driving simulator, eye-tracking, and on-road driving) suggests that while there are subtle differences between the methods, they do not obscure the picture of drivers' risk perception that was produced. Participants consistently under-rated the risk associated with roadside hazards regardless of the mode of stimulus presentation; even though participants could reasonably have been expected to spend more time assessing the roadsides when looking at the still photos (where their speed was zero and there are greater opportunities to look at hazards on the roadside, this did not result in an increase in their ratings of risk for those situations. Similarly, the consistency checks built into the driving simulator testing indicated that the participants' ratings were extremely reliable and reflected the perceptions of risk arising from the properties of the roads rather than random variation. Indeed, the present study represents the first time these different methods of testing have been directly compared using the same stimuli and rating scale.

The use of multiple testing methods not only allowed us to assess the reliability of the findings, they also provided a more complete picture of drivers' risk perception than has been previously possible. The finding that pupil size and the number of saccades increased as the participants' perceived risk increased, suggests that these measures might be used as a useful addition to other measures of perceived risk. Further investigation of these measures is warranted, however, in the light of recent evidence suggesting that pupil size may be related to motor preparation for saccades (Jainta, Vernet, Yand & Kapoula, 2011) and to verify that the increased number of saccades associated with high risk ratings were not unique to the stimuli used in the present experiment.

There are several safety implications of the present research that are worth considering. For example, the finding that narrow shoulders, poles, and ditches were not included in the participants' considerations of risk suggests that investment in eliminating or reducing these roadside hazards should be given a high priority. In contrast, intersections were noticed and considered by the participants, but their risk was still under-rated compared to RPS risk estimates. This suggests that exploring how to make the risks associated with intersections more apparent to drivers, or what aspects of intersections contribute to drivers' risk perceptions would be a valuable contribution to road safety. As noted by earlier researchers, roads with low perceived risk may actually contribute to higher levels of objective risk because drivers are not taking due care at these locations.

Acknowledgements

This research was funded by the New Zealand Automobile Association Research Foundation. The authors would like to acknowledge the patience and hard work of the participants, research assistants, and technical support staff who worked with us on this project. Without their dedication this work would simply not have been possible.

References

- Armsby, P., Boyle, A. J. and Wright, C. C. (1989). Methods for assessing drivers' perception of specific hazards on the road. *Accident Analysis and Prevention*, 21, 45-60.
- Benda, H. V., & Hoyos, C. G. (1983). Estimating hazards in traffic situations. *Accident Analysis and Prevention*, 15, 1-9.
- Brookhuis, K. A., de Waard, D. (2010). Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis & Prevention*, 42, 893–903.
- Burdett, B. R. D., & Nicholson, A. J. (2010). *Speed Management on Rural Roads*. In Proceedings of IPENZ Transportation Group Conference. Christchurch: Institution of Professional Engineers New Zealand.
- Charlton, S. G. (2006). Conspicuity, memorability, comprehension, and priming in road hazard warning signs. *Accident Analysis and Prevention*, 38, 496-506.
- Charlton, S. G. (2007). The role of attention in horizontal curves: A comparison of advance warning, delineation, and road marking treatments. *Accident Analysis and Prevention*, 39, 873-885.
- Charlton, S. G. and Starkey, N. J. (2011). Driving without awareness: The effects of practice and automaticity on attention and driving. *Transportation Research Part F: Traffic Psychology and Behaviour*, 14, 456-471.
- Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis and Prevention*, 37, 461-472.
- Gibson, J. J, Crooks, L. E. (1938). A theoretical field-analysis of automobile-driving. *The American Journal of Psychology*, 11, 453-471.
- Groeger, J. A., & Chapman, P. R. (1996). Judgment of traffic scenes: The role of danger and difficulty. *Applied Cognitive Psychology*, 10, 349–364.
- Jainta, S., Vernet, M., Yang, Q. & Kapoula, Z. (2011). The pupil reflects motor preparation for saccades- even before the eye starts to move. *Frontiers in Human Neuroscience*, 5, 1-10.
- Kanellaidis, G. (1995). Factors affecting drivers' choice of speed on roadway curves. *Journal of Safety Research*, 26, 49-56.
- Kanellaidis, G., & Dimitropoulos, I. (1994). Subjective and objective evaluation of risk on roadway curves. *Traffic Engineering and Control*, 55, 451–454.
- Kanellaidis, G., Zervas, A., & Karagioules, V. (2000). Drivers' risk perception of road design elements. *Transportation Human Factors*, 2, 39-48.
- Lewis Evans, B. and Charlton, S. G. (2006). Explicit and implicit processes in behavioural adaptation to road width. *Accident Analysis and Prevention*, 38, 610-617.
- Liston, B.D., Krukowski, A.E., & Stone, L.S. (2013). Saccade detection during smooth tracking. *Displays*, 34, 171-176.

- Lund, A. K., & O'Neill, B. (1986). Perceived risks and driving behaviour. *Accident Analysis and Prevention*, 18, 367–370.
- Näätänen, R., and Summala, H. (1974). A model for the role of motivational factors in drivers' decision-making. *Accident Analysis & Prevention*, 6, 243-261.
- Pelz, D. C. and Krupat, E. (1974). Caution profile and driving record of undergraduate males. *Accident Analysis and Prevention*, 6, 45–58.
- Taylor, D. H. (1964). Drivers' galvanic skin response and the risk of accident. *Ergonomics*, 7, 439-451.
- Waibl, G., Tate, F., and Brodie, C. (2012). The development of a proactive road safety assessment tool – KiwiRAP. In *Proceedings of Australasian Road Safety Research, Policing and Education Conference 2012*, 4 - 6 October 2012, Wellington, New Zealand, (pp 1-11), Mawson, ACT: Australasian College of Road Safety.
- Watts, G. R., and Quimby, A. R. (1980). *Aspects of Road Layout That Affect Drivers' Perception and Risk Taking*. TRRL Report 920. Crowthorne, Berkshire, England: Transport and Road Research Laboratory.
- Wilde, G. J. S. (1982). The theory of RHT: Implications for safety and health. *Risk Analysis*, 2, 209-226.