PERGAMON

Transportation Research Part F 4 (2001) 201–212

TRANSPORTATION RESEARCH PART F

www.elsevier.com/locate/trf

Risk perception in novice drivers: the relationship between questionnaire measures and response latency

Paul Farrand [a], Frank P. McKenna [b,*]

[a] *Psychology Section, Department of Human Science and Medical Ethics, St. Bartholomew's and the Royal London School of Medicine and Dentistry, Turner Street, London E1 1AD, UK*

[b] *Department of Psychology, University of Reading, Whiteknights, Reading RG6 6AL, UK*

Received 31 July 2000; received in revised form 27 July 2001; accepted 8 August 2001

Abstract

The extent to which questionnaire based measures of driving risk, driving ability and accident likelihood are associated with response latency based measures obtained on a hazard perception test was examined. In Experiment 1 questionnaire evaluations of driving in general were obtained and correlated with hazard perception performance. In Experiment 2 questionnaire evaluations and hazard perception performance were obtained when drivers viewed the same driving scenes. In neither experiment did questionnaire responses correlate significantly with hazard perception performance. Additionally while in both experiments no difference in hazard perception performance arose between males and females, females rated driving as more risky and their ability to be lower than males. The results indicate independence between questionnaire and response latency measures of hazard perception. However the possibility that both approaches should be adopted within a single framework is raised.

Keywords: Hazard perception; Risk perception; Questionnaire; Response latency

1. Introduction

There is much theoretical and practical interest in the issue of risk perception. The extent to which people can and do make judgements about the hazards they face has been a source of ongoing debate (Pascoe & Pidgeon, 1995). Investigation of the dimensions used to characterise these judgements has shifted from the relatively simple judgements of associated fatalities to judgements involving factors such as controllability and familiarity. In addition, the methods used

* Corresponding author. Tel.: +44-118-931-8523; fax: +44-118-931-6715.
E-mail address: f.p.mckenna@reading.ac.uk (F.P. McKenna).

1369-8478/01/$ - see front matter
PII: S1369-8478(01)00024-9

to investigate these issues are varied including questionnaires and performance measures. For example, in the driving field a range of rating measures (Finn & Bragg, 1986; Groeger & Chapman, 1996; Groeger & Grande, 1996; McKenna, Stanier, & Lewis, 1991; Rutter, Quine, & Albery, 1998) and performance measures (Horswill & McKenna, 1999) have been employed to investigate risk perception. While these different measures have had overlapping concerns the research agendas have not been identical.

Given the high accident rates of young drivers this group has been the focus of many investigations. For example, Jonah (1997) has reviewed the evidence indicating a relationship between sensation seeking and risky driving. It has also been shown that accident involvement has been associated with self-reported driving violations (Parker, Reason, Manstead, & Stradling, 1995). In addition, Finn and Bragg (1986) have noted that it is possible that some young drivers are less inclined to perceive the risks of driving. Another possibly related issue is the way that young drivers perceive their own ability. Following Weinstein (1980) a good deal of work has examined the possibility that drivers may be unrealistically optimistic about their driving ability and accident likelihood. The findings here have been that drivers across a range of ages and cultures tend to have a high opinion of their driving ability (Sivak, 1997; Svenson, 1981). As Groeger and Grande (1996) have noted, the extent to which these self-assessments are related to more objective assessments is an interesting point.

While questionnaires have generally been employed to measure self-assessments of risk and skill, alternative methods have been used for more objective assessments. In the present context the objective assessment considered was hazard perception. In the earliest work into hazard perception drivers viewed film of traffic conditions after which they used a checklist as a basis to select the features they felt were important (Spicer, 1964; as cited by Pelz & Krupat, 1974). The critical novel feature of this work was that drivers were making judgements of real dynamic traffic situations. Supporting the importance of dynamic traffic scenes Cohen (1981) had found that there were important differences between viewing static scenes and dynamic scenes. In particular there were differences in the fixation times and what was fixated, leading Cohen (1981) to conclude that static visual scenes were not appropriate for assessing driving abilities. Pelz and Krupat (1974) extended the hazard perception methodology by requiring drivers to make their judgements online, that is during presentation of the driving scene. In principle, this should increase the ecological validity since online responses are a key feature of real driving. Both the Spicer and the Pelz and Krupat (1974) studies reported that their methods distinguished between accident involved and non-accident involved drivers. Watts and Quimby (1979) expanded the method by devising a simulator in which drivers judged how risky and dangerous they considered a set of traffic conditions. Similar to the method adopted by Pelz and Krupat (1974), drivers made their responses by adjusting a lever to indicate the degree of risk during presentation of the driving situation.

Recent developments in hazard perception have seen a shift away from using a simulator. Instead the vehicle is eliminated and within some paradigms the driver is requested to make a discrete response, usually by a button press, as soon as a hazard is detected (Hull & Christie, 1992; McKenna & Crick, 1991). Hazard perception performance is measured in terms of the time taken to respond to the hazard. However, the methods used to obtain responses from drivers has varied considerably. For example, in some cases drivers watch traffic scenes and continually adjust a lever to indicate the degree of safety (Pelz & Krupat, 1974) or the level of perceived risk (Deery &

Love, 1996; Watts & Quimby, 1979; West, Wilding, French, Kemp, & Irving, 1993). In these instances hazard perception performance is measured in terms of the time taken to respond to the hazard but in this case response latency is measured in a more complex manner. Response latency represents the difference in time between the onset of the hazard and a rapid increase in the level of perceived risk as indicated by adjustments to the lever. It is possible that different latency based hazard perception tests are not identical. For example, while the test used by Chapman and Underwood (1998) was not able to distinguish between novice and experienced drivers the test developed by McKenna (McKenna & Crick, 1991; McKenna, Waylen, & Burkes, 1998) does differentiate drivers on the basis of experience.

In other paradigms the focus has not been so much upon taking response latency as an index of risk perception but rather judgements across a range of rating scales are considered. For example, some studies have asked drivers to rate the degree of danger in a given scene (Kruysse & Wijlhuizen, 1992; Ogawa, Renge, & Nagayama, 1996). While other studies have required drivers to complete a short questionnaire which has included judgements of level of risk, degree of concentration required and level of stress (Groeger & Chapman, 1996).

Given the range of measures that have been employed under the broad umbrella of risk perception an issue arises as to the extent of overlap, if any, among the measures. For example, do those who rate driving as very risky have a high or low estimate of their ability to deal with the risks of driving? One might argue that those who rate driving as very low in risk do so because they have a high estimate of their ability to deal with the risks of driving. Alternatively, it might be that it is only those who have a high estimate of their ability to deal with the risks of driving who appreciate that driving is very risky. Similarly an issue arises as to the relationship between a more objective measure of risk perception, in this case hazard perception, and more subjective measures of risk perception as measured by rating scales.

There is some evidence to suggest that rating scale measures of risk perception may be tapping a different aspect of risk perception. For example, Ogawa et al. (1996) and Soliday (1975) both report that women perceive the risks involved in driving to be greater than men. However, in considering more objective measures of risk perception it has been found that there is no difference between men and women in their ability to detect hazards (McKenna et al., 1998). While there may be a number of different reasons for this discrepancy the role of driver experience should be noted. It is generally known that women drive less than men. It is possible then that the higher rating of risk among women is related to their relative inexperience. Consistent with the driver experience account of this discrepancy is the fact that the McKenna et al. (1998) study which showed no sex difference employed groups of men and women that were matched in driver experience.

2. Experiment 1

Overall, the aim of this paper was to consider the relationship between a number of rating scale measures of risk perception and a particular performance based measure of risk perception, namely the hazard perception test. Is it the case that rating scale and latency based measures are assessing the same characteristic? An additional method of exploring the same issue was to consider sex differences in risk perception. In particular, is there any evidence of a dissociation between rating scales and performance measures. Evidence of the latter would occur if it were possible to observe simultaneously a sex difference in rating scales in the absence of any sex difference in hazard perception.

2.1. Method

Participants: 133 participants (61 males, 72 females) with a mean age of 18.6 years volunteered to take part in the experiment and were paid an honorarium. All participants had passed the British Driving Test, had a maximum of 3 years post-test driving experience and were students at the University of Reading or at local sixth form schools or colleges. Male drivers had been driving for an average of 20.4 months (S.D. 9.1) with an average weekly mileage of 59.5 miles (S.D. 67.8). Women drivers had been driving for an average of 22.2 months (S.D. 9.6) with an average weekly mileage of 59.2 miles (S.D. 65.7). There were no significant differences between the males and females in terms of months driving or weekly mileage (comparisons $p > .05$).

2.2. Procedure

Participants were initially required to complete three 11-point Likert scale based questions. An 11-point scale was chosen in order to allow participants the opportunity to provide a relatively fine discrimination in their ratings. The first question – 'Risk', was concerned with the participants perceptions of the risks involved in driving (How would you estimate the risks involved in driving?) where 1 represented 'Very Risky' and 11 'Not Risky'. The second question – 'Accident Likelihood', examined perceptions of involvement in a driving accident (Compared to the average driver how likely do you think you are of being involved in a driving accident?) where 1 represented 'Much Less Likely' and 11 'Much More Likely'. The final question – 'Driving Ability', examined perceptions of driving ability (Compared to the average driver how capable do you think you are of dealing safely with the risks involved in driving?) with 1 representing 'Much Less Capable' and 11 'Much More Capable'. For all three questions the value of 6 was indicated as being 'Average'.

After completing the questionnaires participants were then required to complete the hazard perception test which is a videotaped sequence of 10 road-based scenarios (8 hazardous, 2 non-hazardous) lasting approximately 10 min. The hazard perception test was presented using a JVC TM-210 colour video monitor (40 cm $\times$ 30 cm) and JVC BR-S6000E stereo video recorder, which were interfaced with a Phillips PC-590 computer and response button. The nature of the hazards varied across scenario and the hazards were developed on the basis of information obtained from various expert driving groups such as Police drivers and driving instructors. Informal interviews and a form of the Critical Incident technique were used (Flanagan, 1954). On the basis of such information a series of staged and naturally occurring hazardous scenarios were videotaped while driving around a few major British towns and surrounding countryside. Examples of scenes included the following: while proceeding down an urban road a vehicle approaches an oncoming junction at speed; a pedestrian steps out on to the road; an oncoming vehicle overtakes a parked car and is on a collision course.

To ensure that the scenes would look as close as possible to those which would be viewed while driving the video camera was mounted as closely as possible to the driver to provide a driver's-eye view of the road situation. Participants were instructed to respond as fast as possible by pressing a response button whenever they detected any hazard. A hazard was defined as any event that could develop into an accident risk. The computer recorded response latencies. To ensure that it was possible to detect all hazards the clock was started before the hazard was on screen for eight out of the ten hazards. For the other two the start time was determined empirically by starting the clock earlier than the fastest responses from an expert group of drivers. Three practice trials (two hazards, one non-hazard) were given prior to the start of the test. At the end of the experiment participants were asked to complete a brief questionnaire concerned with general demographic variables including age, sex and driving experience.

2.3. Results

Independent samples *t*-tests were performed on the hazard perception data with sex (male vs. female) as the independent variable and hazard perception score and misses as the dependent variables. Since the mean and the variance of the latencies to the hazards varied across scenarios these were converted to a standardised score in order to produce an overall hazard perception score. This allowed each hazard to contribute equally to the overall hazard perception score. Using a prior sample of over 400 hundred drivers who varied in age and driving experience a sigma scale was constructed. This scale has a mean of 50 and a S.D. of 16.66. Each response latency from the present study was converted to this scale in such a way that high scores reflected better hazard perception performance. In scenarios where no response was made to a pre-defined hazard a miss was recorded. Overall scores were computed by averaging the standardised scores.

As can be seen in Table 1 differences in hazard perception performance ($t(131) = .41$; $p > .05$) and mean number of misses ($t(131) = .78$; $p > .05$) between males and females were both non-significant.

The Likert risk ratings were approximately normally distributed and more formally the Shapiro and Francia (1972) test of normality was carried out. No significant departure from normality was detected for the risk ratings. Females rated driving to be significantly more hazardous than males ($t(131) = 10.0$, $p < .001$). In addition females rated themselves to be significantly more likely to be involved in an accident ($t(131) = 3.9$, $p < .001$). Finally, in rating their ability to deal with the risks of driving females rated themselves as significantly worse than males ($t(131) = 5.5$, $p < .0001$).

Pearson Product–Moment correlation coefficients were performed to examine the relationship between hazard perception performance and participant ratings to the three assessment questions.

Table 1
Male and female mean (standard deviations) sigma hazard perception score, misses and responses to ‘risk’, ‘driving ability’ and ‘accident likelihood’ questions applied to driving in general

	Male	Female
Sigma hazard perception score	38.5 (17.2)	39.8 (17.8)
Misses	1.8 (1.2)	1.8 (1.2)
Risk	7.3 (1.6)	4.4 (1.7)
Driving ability	6.7 (1.5)	5.2 (1.5)
Accident likelihood	5.0 (1.8)	6.3 (2.0)

Hazard perception was not found to be significantly correlated with the assessment of risk ($r = -.08$; $n = 133$, $p > .05$), driving ability ($r = -.15$; $n = 133$, $p > .05$) or accident likelihood ($r = .09$; $n = 133$, $p > .05$). This pattern also emerged when the ratings of males and females to the three questions were separately correlated with hazard perception performance. Correlations for males and females, respectively, with risk ($r = .04$, $n = 61$; $r = -.16$, $n = 72$), driving ability ($r = -.16$, $n = 61$; $r = -.13$, $n = 72$) and accident likelihood ($r = .06$, $n = 61$; $r = .11$, $n = 72$) were all non-significant (all $p > .05$).

Separate correlations performed between the ratings for the three questions indicated that answers to all the questions were significantly correlated with each other, risk – driving ability ($r = .41$, $n = 133$, $p < .0001$), risk -accident likelihood ($r = -.40$, $n = 133$, $p < .0001$), driving ability – accident likelihood ($r = -.45$, $n = 133$, $p < .0001$).

2.4. Discussion

The results from Experiment 1 indicate that there is little relationship between levels of risk perception in general and hazard perception in particular. Adding to this conclusion is the sex difference dissociation between rating scale measures and performance based measures of risk perception. While there was no significant difference in hazard perception response latency between male and female drivers, a finding consistent with previous research (McKenna et al., 1998) females rated the risks involved in driving as greater than males, consistent with research by Ogawa et al. (1996) and Soliday (1975). The possibility that these different results might be accounted for by varying patterns of driver experience is discounted in the present experiment since the dissociation is observed in a single sample who responded to both the hazard perception latency test and the rating scale measures.

It is interesting to note that while there is little relationship between hazard perception and the rating scale measures there are significant relationships between the rating scale measures. Those who rate their ability to deal with the risks of driving as less also rate their accident likelihood as greater. This relationship is not inevitable since it is possible to perceive that one has a high ability to deal with the risks of driving while also having a higher accident likelihood as more risks may be taken. For example, Williams and O'Neill (1974) found that those who would be rated as having greater driving skills also had a higher number of accidents. Additionally the rating scale measures revealed that those who rate their personal accident likelihood as greater also rate the risks of driving in general as greater. Finally, those who have a high rating of their ability to deal with the risks of driving also rate driving as less risky in general. Once again this relationship is not inevitable as it is readily possible to conceive of a scenario in which those who have a high estimate of the risks of driving in general also have a high estimate of their personal abilities to cope with these risks.

3. Experiment 2

Experiment 1 found no evidence of a relationship between general ratings of perceptions of risk and hazard perception performance. Experiment 2 was designed to take this issue a stage further by assessing risk ratings and hazard perception when drivers were viewing the same driving scenes.

In a large-scale study of hazard perception Ogawa et al. (1996) used a measure that was based on ratings of driving scenes. Having observed driving scenes filmed from the perspective of the driver participants rated the degree of hazard in each of the scenes. It was found that, overall, women rated the scenes as more hazardous than the men. One aim of Experiment 2 was to determine whether hazard perception as measured through a rating scale method would provide the same pattern of results as that produced through a latency measure. Since the term hazard perception is being used to refer to both rating scale measures and latency based measures it is important to determine if they are measuring the same factor. Drivers were asked to perform the hazard perception test and when they pressed the button to indicate they had detected a hazard they were then asked to rate this hazard.

3.1. Method

Participants: 70 participants (30 males, 40 females) with a mean age of 20.3 years volunteered to take part in the experiment and were paid an honorarium. All participants had passed the British Driving Test, had a maximum of 3 years post-test driving experience and were students at the University of Reading or at local sixth form schools or colleges. Male drivers had been driving for an average of 22.4 months (S.D. 9.5) with an average weekly mileage of 76.6 miles (S.D. 130.6). Women drivers had been driving for an average of 21.5 months (S.D. 9.4) with an average weekly mileage of 67.2 miles (S.D. 67.3). There were no significant differences between the males and females in terms of months driving or weekly mileage ($p > .05$). None of the participants had previously taken part in Experiment 1.

3.2. Procedure

The hazard perception test used in Experiment 1 was adapted for use in Experiment 2. The procedure was the same with the exception that every time the response button was pressed the video was stopped and participants were required to complete two rating scales. To maintain compatibility with Experiment 1 the same rating scales were used – ‘Risk’ and ‘Driving Ability’. First drivers were asked “How would you estimate the risks involved in driving?” though in this case they were asked to estimate the risk in the scene they had just observed and responded to. Second, they were asked “Compared to the average driver how capable do you think you are of dealing safely with the risks involved in driving?” and once again it was made clear that they should make their judgement concerning the driving event they had just observed and responded to. In this way for each hazard there was a latency and associated with this response there were two rating scale measures. The ‘Accident Likelihood’ question from Experiment 1 was not included for use as it was felt that this question was less easily applied to the specific hazardous scenarios but more suited to driving risks in general. At the end of the experiment participants completed a brief questionnaire concerned with general demographic variables.

3.3. Results

As with Experiment 1, prior to analysis hazard perception response latency data were converted into a standardised sigma scale with each sigma point inverted. Participants were given a score for the hazard perception test by collapsing their sigma points across the scenarios in each test. In scenarios where no response was made to a pre-defined hazard within a time window a miss was recorded and no value placed in the analysis.

Independent samples *t*-tests were performed on the hazard perception data with sex (male vs. female) as the independent variable and standardised sigma hazard perception score and misses as the dependent variables. As can be seen in Table 2 there were no significant differences in hazard perception performance ($t(68) = .8$; $p > .05$). Additionally the mean number of misses did not differ significantly ($t(68) = .67$; $p > .05$).

Females rated the level of risk across the scenarios as significantly higher than males ($t(68) = 4.2$, $p < .001$). In addition, females rated their ability to deal with the risks presented in the scenarios as significantly less than males ($t(68) = 2.3$, $p < .05$).

In order to address the issue of whether there was any relationship between hazard perception scores and risk ratings three different analyses were employed. First the average hazard perception scores were computed and the average risk ratings were computed and then the Pearson Product–Moment correlation coefficients were calculated. The Likert risk ratings were approximately normally distributed and more formally the Shapiro and Francia (1972) test of normality was carried out. No significant departure from normality was detected for the average risk ratings. The Pearson Product–Moment correlation coefficients were as follows: neither the assessment of risk ($r = .10$, $n = 70$, $p > .05$) nor perception of ability to deal with the risks ($r = -.10$, $n = 70$, $p > .05$) correlated significantly with hazard perception latency. However, the correlation between the two risk perception ratings was significant ($r = .69$, $n = 70$, $p < .01$) indicating that those who had a high estimate of their ability to deal with the risks presented in the scenarios judged that the scenarios were less risky.

The method just described might be called correlating averages in that the averages for each of the measures was calculated and then the correlations carried out. An alternative advocated by Dunlap, Jones, and Bittner (1983) is to average the correlations, that is to calculate the correlations first and then take an average of the correlations. This was done in two ways. First, for each participant the correlation coefficients were calculated and then these were averaged. A problem in averaging correlation coefficients is that the sampling distribution is increasingly skewed for high correlations. The standard solution to this problem, which was employed in the present study, is to use Fisher's Z transformation prior to averaging. We will first consider the relationship between hazard perception and risk assessment. Of the 70 participants only five participants showed a significant relationship between hazard perception latency and risk assessment. Three participants showed a significant positive relationship and two a significant negative

Table 2
Male and female mean (standard deviations) sigma hazard perception score, misses and responses to 'risk' and 'driving ability' questions applied to individual hazardous scenes observed and responded to

	Male	Female
Sigma hazard perception score	38.9 (9.2)	37.2 (7.9)
Misses	1.9 (1.3)	2.1 (1.5)
Risk	5.3 (1.1)	3.9 (1.5)
Driving ability	4.9 (1.4)	4.2 (1.3)

relationship. The average correlation between hazard perception latency and risk assessment was −.026 indicating that there was little shared variance. In considering the relationship between hazard perception latency and the rating of ability to deal with the risk only two people showed a significant relationship. The average correlation was −.005 once again indicating little shared variance. In considering the relationship between the two risk perception ratings 19 participants revealed a significant relationship, all in the same direction, indicating that a high estimate of ability to deal with the risks was associated with a judgement that the scenario was less risky. The average correlation between the two risk ratings was .62 indicating considerable shared variance.

The final method used to determine the relationship between hazard perception latency and risk ratings was to calculate the correlations separately for each of the hazard perception scenarios and then average the correlations across scenarios. Once again Fisher's Z transformation was applied prior to averaging. In considering the relationship between hazard perception latency and risk assessment it was found that there was no significant relationship for any of the scenarios. The average correlation was −.083 indicating that there was little shared variance. In considering the relationship between hazard perception latency and the rating of ability to deal with the risk it was found that there was no significant correlation for any of the scenarios. The average correlation was .0123 indicating that there was little shared variance. In considering the relationship between the two risk perception ratings it was found that the correlation was significant for each and every scenario indicating that a high estimate of ability to deal with the risks was associated with a judgement that the scenario was less risky. The average correlation was .565 indicating considerable shared variance.

3.4. Discussion

Once again there is little evidence of a relationship between ratings of risk perception and a response latency measure. While the term hazard perception is used to describe each of these measures it is clear that different attributes are being assessed. This conclusion is reinforced by the pattern of sex differences. Even when responding to exactly the same driving scenarios there was no difference between men and women in their response latencies but women estimated that the risks were greater than men while also having a lower estimate of their ability to deal with the risks.

4. General discussion

The pattern of results obtained in this paper go some way towards clarifying the relationship between different measures of risk perception in general and the different paradigms used to measure hazard perception in particular. It seems evident that paradigms designed to measure hazard perception that are based upon response latency (Hull & Christie, 1992; McKenna & Crick, 1991) measure different processes to those using rating scales (Ogawa et al., 1996). The results of Experiment 1 clearly indicate that a significant relationship does not exist between hazard perception response latency and questionnaire measures ascertaining general levels of driving risk, driving ability or accident likelihood. Experiment 2 further supports these findings by replicating the same pattern of results when all judgements are constrained by forcing responses to be made on the same driving scenarios.

That response latency and rating scales do not measure the same aspects of hazard perception is further supported by the dissociation in the pattern of sex differences. While both experiments reported no difference in response latency between males and females, differences did exist in responses to the rating scale measures. Consistent across both experiments females rated driving to be more hazardous and their ability to cope with the risks as less than males, a finding consistent with previous research (Ogawa et al., 1996; Soliday, 1975). It is possible that these particular sex differences are part of the more general difference in risk perception that has previously been reported. In other words it has been found that across a broad range of hazards women rate risks as greater (Savage, 1993).

The present investigation addresses one potential limitation that is present in much work on rating scales. It is possible that in asking people to rate such broad categories of behaviour as driving that the sex difference emerges in the specific scenarios individuals bring to mind prior to judgement. In other words it is possible that men and women construct different driving scenarios and that the difference that emerges is due to the fact that women construct more dangerous scenarios. It should be noted that most questionnaires provide very general categories with little further guidance, leaving participants with considerable latitude in what they decide to rate. Experiment 2 eliminated this possibility by providing men and women with exactly the same specific concrete scenarios to rate. It was notable that even under these circumstances the sex difference still emerged.

In Section 1 it was recognised that the pattern of sex differences might have been due to driver experience. In other words men, on average, have a higher driving mileage than women. Differences observed in rating scale measures might, therefore, be due to either driver experience or gender. The fact that there were no differences between men and women on a response latency measure of hazard perception (McKenna et al., 1998) could be interpreted as indicating that the latency measure was tapping a different process to the rating scale measures. Alternatively it might be noted that the McKenna et al. (1998) study employed men and women who were matched in driving experience. It is possible, therefore, that the studies which show differences in risk ratings employ samples of men and women who differ in driving experience while those studies which show no difference in response latency employ samples that are matched in driving experience. The present work demonstrates that the dissociation is not due to a sampling issue since it is observed when the same sample perform latency and rating scale measures.

While there is little support for the notion that the different measures assess the same aspect of driving behaviour it is possible that each measure makes a useful, if independent, contribution to our understanding. For example, it seems likely that response latency measures will make their contribution through what Rumar (1990) has labelled the basic driver error, namely, late detection. In essence he argues that delayed detection of hazards is a major factor in road traffic collisions. Latency measures provide a means for assessing individual differences in this ability and for assessing factors that might impact on this process. An example of the latter would be that this process does not appear to be an automatic one and can be disrupted by other secondary tasks (McKenna & Farrand, 1999).

Questionnaire measures of driving risk, ability, and accident likelihood may be more useful in gaining an understanding of the strategic components that can influence behaviour. Such strategic factors might take the form of general attitudes surrounding ability that can in turn influence whether drivers expose themselves to hazardous driving scenarios. A well replicated finding is that young drivers perceive that relative to the average driver they themselves are more skilled (Svenson, 1981) and less likely to be involved in an accident (McKenna, 1993). Weinstein and Lyon (1999) have argued that optimistic biases about risk are a barrier to the adoption of risk-reducing behaviour. In theory these same perceptions may not only be a barrier to the adoption of safer behaviour they may also be part of the process that facilitates risk-taking behaviour. It has already been noted that men relative to women have a higher estimate of their ability and a lower estimate of the risks in driving. The extent to which these perceptions are related to the findings that males are much more likely to engage in risk taking behaviour (Byrnes, Miller, & Schafer, 1999) is not yet clear.

It should be fairly clear that while the research does not support the notion that questionnaire and response latency measures of risk perception tap the same process this should not imply that one approach should be adopted at the expense of the other. It is possible that both approaches contribute significantly to our understanding of separate mechanisms involved in risk perception, and that each approach measures a qualitatively different process. Indeed outside of the domain of driving behaviour an attempt at developing a model of pilot performance has progressed while maintaining a wide range in the type and function of measurement. The situational awareness model (see, for example, Endsley, 1995) employs verbal reports, questionnaires and performance based methodologies in a model which attempts to understand both the dynamic interaction of elements and the generation of future states to be engaged in the environment.

Similarly it would appear prudent that the future direction of driving research should be to work towards developing a model of risk perception based upon both questionnaire and response latency paradigms. This model would therefore attempt to fully understand not just how the driver currently perceives their environment but also how this influences their ability to anticipate and deal with future events.

References

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk-taking: a meta-analysis. *Psychological Bulletin*, *125*, 367–383.

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: danger and experience. *Perception*, *27*, 951–964.

Cohen, A. S. (1981). Car drivers' pattern of eye fixations on the road and in the laboratory. *Perceptual and Motor Skills*, *52*, 515–522.

Dunlap, W. P., Jones, M. B., & Bittner, A. C. (1983). Average correlations vs. correlated averages. *Bulletin of the Psychonomic Society*, *21*, 213–216.

Deery, H. A., & Love, A. W. (1996). The effect of a moderate dose of alcohol on the traffic hazard perception profile of young drink-drivers. *Addiction*, *91*, 815–827.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, *37*, 32–64.

Finn, P., & Bragg, B. W. E. (1986). Perception of the risk of an accident by young and older drivers. *Accident Analysis and Prevention*, *18*, 289–298.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*, 327–358.

Groeger, J. A., & Chapman, P. R. (1996). Judgement of traffic scenes: the role of danger and difficulty. *Applied Cognitive Psychology*, *10*, 349–364.

Groeger, J. A., & Grande, G. E. (1996). Self-preserving assessments of skill? *British Journal of Psychology*, *87*, 61–79.

Horswill, M. S., & McKenna, F. P. (1999). The effect of interference on dynamic risk-taking judgements. *British Journal of Psychology*, *90*, 189–199.

Hull, M., & Christie, R. (1992). Hazard perception test: the geelong trial and future development. *National road safety seminar*, *vol. 2*. Wellington, New Zealand: The Official Committee on Road Safety and Road Traffic Safety Research Council.

Jonah, B. A. (1997). Sensation seeking and risky driving: a review and synthesis of the literature. *Accident Analysis and Prevention*, *29*, 651–665.

Kruysse, H. W., & Wijlhuizen, G. J. (1992). Why are experts not better in judging the danger of filmed traffic conflicts. *Accident Analysis and Prevention*, *24*, 227–235.

McKenna, F. P. (1993). It won't happen to me: unrealistic optimism or illusion of control. *British Journal of Psychology*, *84*, 39–50.

McKenna, F. P., & Crick, J. L. (1991). Experience and expertise in hazard perception. In G. B. Grayson, & J. F. Lester (Eds.), *Behavioural research in road safety*. Crowthorne, UK: Transport and Road Research Laboratory.

McKenna, F. P., & Farrand, P. (1999). The role of automaticity in hazard perception. In G. B. Grayson, & J. F. Lester (Eds.), *Behavioural research in road safety*. Crowthorne, UK: Transport and Road Research Laboratory.

McKenna, F. P., Stanier, R. A., & Lewis, C. (1991). Factors underlying illusory self-assessment of driving skill in males and females. *Accident Analysis and Prevention*, *23*, 205–208.

McKenna, F. P., Waylen, A., & Burkes, M. E. (1998). Male and female drivers: how different are they? AA Foundation for Road Safety Research.

Ogawa, K., Renge, K., & Nagayama, Y. (1996). *The structure of the perception of hazards in driving scenes*. Osaka, Japan: Osaka University.

Parker, D., Reason, J. T., Manstead, A. S. R., & Stradling, S. G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics*, *38*, 1036–1048.

Pascoe, E., & Pidgeon, N. (1995). Risk orientation in dynamic decision making. In J. P. Caverni, & M. Bar-Hillel (Eds.), *Contributions to decision making*. Amsterdam: Elsevier.

Pelz, D. C., & Krupat, E. (1974). Caution profile and driving record of undergraduate males. *Accident Analysis and Prevention*, *6*, 45–58.

Rumar, K. (1990). The basic driver error: late detection. *Ergonomics*, *33*, 1281–1290.

Rutter, D. R, Quine, L., & Albery, I. P. (1998). Perceptions of risk in motorcyclists: unrealistic optimism, relative realism and predictions of behaviour. *British Journal of Psychology*, *89*, 681–696.

Savage, I. (1993). Demographic influences on risk perceptions. *Risk Analysis*, *13*, 413–420.

Shapiro, S. S., & Francia, R. S. (1972). An approximate analysis of variance test for normality. *Journal of the American Statistical Association*, *67*, 215–216.

Sivak, M. (1997). Recent psychological literature on driving behaviour: what, where and by whom? *Applied Psychology: An International Review*, *46*, 303–310.

Soliday, S. M. (1975). Development and preliminary testing of a driving hazard questionnaire. *Perceptual and Motor Skills*, *41*, 763–770.

Svenson, O. (1981). Are we all less risky and more skilful than our fellow drivers? *Acta Psychologica*, *47*, 143–148.

Watts, G. R., & Quimby, A. R. (1979). Design and validation of a driving simulator for use in perceptual studies. *TRRL report LR907*. Crowthorne, UK: Transport and Road Research Laboratory.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology*, *39*, 806–820.

Weinstein, N. D., & Lyon, J. E. (1999). Mindset, optimistic bias about personal risk and health-protective behaviour. *British Journal of Health Psychology*, *4*, 289–300.

West, R., Wilding, J., French, D., Kemp, R., & Irving, A. (1993). Effect of low and moderate doses of alcohol on driving hazard perception latency and driving speed. *Addiction*, *88*, 527–532.

Williams, A. F., & O'Neill, B. (1974). On-the-road driving records of licensed race drivers. *Accident Analysis and Prevention*, *6*, 263–270.