OVERESTIMATION OF SKILLS AFFECTS DRIVERS' ADAPTATION TO TASK DEMANDS

Saskia de Craen,[1] Divera A.M. Twisk,[1] Marjan P. Hagenzieker,[1]
Henk Elffers,[2] and Karel A. Brookhuis[3]
[1]SWOV, Institute for Road Safety Research
Leidschendam, the Netherlands
E-mail: Saskia.de.Craen@SWOV.nl
[2]NSCR, Institute for the Study of Crime and Law Enforcement
Leiden, the Netherlands
[3]Delft University of Technology
Delft, the Netherlands

Summary: Inadequate self-assessment, and specifically, overestimation of skill, results in insufficient adaptation to task demands, which can manifest itself on different levels of the driving task. A total of 130 drivers (83 novice and 47 experienced drivers) participated in an on-road driving assessment. Their performance in this assessment (i.e., fail or pass) was compared to the participants' reported confidence in their driving skills (i.e., high or low confidence), resulting in three calibration groups: a) well-calibrated drivers (reported confidence matched performance on assessment), b) overconfident drivers (high confidence but failed assessment) and c) insecure drivers (low confidence but passed assessment). Furthermore, participants completed a questionnaire which focused on choices made on the strategic and manoeuvring level of the driving task. No significant difference was found between the calibration groups for the strategic level. Overconfident drivers reported significantly more violating behaviour than the well-calibrated and the insecure drivers. At the manoeuvring level, overconfident drivers showed significantly less instances of adaptation to traffic complexity. In conclusion, the current study suggests that overconfidence is related to inadequate adaptation to task demands.

INTRODUCTION

For safe driving, a driver's capabilities have to match (or exceed) the task demands of a traffic situation (Fuller, 2005). A driver can influence the tasks demands by, for example, adjusting speed or headway. It is assumed that the decision to increase or decrease task demands is influenced by the driver's assessment of the complexity of the traffic situation on the one hand, and of his driver ability on the other hand. Balancing task demands and capabilities has been named "calibration." It is postulated that novice drivers are particularly poor at calibration and that improvement in calibration explains the steady decrease in crash rates in the first months after licensing (Brown & Groeger, 1988; Gregersen, 1995; Kuiken & Twisk, 2001).

The present study mainly deals with the influence of self-assessment of skills on the adaptation to task demands. Inadequate self-assessment of driving skills can go either way, underestimation or overestimation of skills. Most research has been focused on overestimation of driving skills, as that is considered the most dangerous form of inadequate self-assessment (Gregersen, 1995; Mathews & Moran, 1986). The current study, however, also addresses underestimation of skill.

The hypothesis of this study is that inadequate self-assessment of skills results in an insufficient adaptation to the task demands, which can manifest itself on different levels of the driving task (e.g., Fuller & Santos, 2002; Rasmussen, 1986). In the current study, we distinguish adaptation on the two highest levels of Michon's (1984) driving task hierarchy. First, on the *strategic* level, the driver might avoid a difficult junction or choose (not) to drive in the dark. Second, on the *manoeuvring* level, a driver might not sufficiently reduce speed or increase headway when encountering more complex situations. The influence of inadequate self-assessment on reported violating behaviour is also studied, which can be viewed as both strategic, as it involves a strategic choice (not) to violate traffic regulations, and manoeuvring, as the specific situation has a large influence on driver behaviour (Reason, 1990). The lowest level of the driving task hierarchy (i.e., the control or vehicle handling level) was not considered here, as adaptation to task demands encompasses a conscious assessment to be made.

METHOD

Participants

The total sample consisted of novice (n = 83, 52% male) and experienced drivers (n = 47, 57% male). The novice group had 6 months driving experience (Mean age = 20; SD = 1.7) and the experienced group at least 10 years (Mean age = 41; SD = 5.1).

Instruments

Questionnaire. To provide an indication of reported driving behaviour at the strategic and manoeuvring level of the driving task and driver confidence, an online questionnaire was administered. In addition to some background information (e.g., age, occupation), the questionnaire included a five-point scale concerning driver confidence (very confident to not at all) and questions regarding the adaptation to task demands. Relating to the strategic level, drivers were asked: a) whether, in the last 4 months, they had cancelled an intended driving trip due to adverse conditions (e.g., darkness or weather conditions); and b) whether, in the previous 4 months, they had avoided a difficult intersection.

To measure adaptation to task demands on the manoeuvring level, the Adaptation Test was applied (De Craen, Twisk, Hagenzieker, Elffers, & Brookhuis, 2007). This test consists of several traffic situations presented in two almost identical photographs. The photographs differ in one single detail relevant to the driving task, increasing the complexity of the situation (e.g., the presence or absence of pedestrians on the pavement, see Figure 1). The photographs are presented in random order, and the complex and simple photographs of one situation never directly succeed one another. Respondents assess at what speed they *would* drive in the depicted situation. A response is considered 'correct' when the reported speed is lower in the complex situation than in the corresponding simple situation. The adaptation score is the sum of all correct responses (1 credit point per correct response; no negative points are assigned for incorrect responses). An evaluation of the Adaptation Test indicated that it is a valid instrument to measure adaptation of reported driving speed to a specific situation (De Craen et al., 2007).

Figure 1. Example of a simple (left) and complex (right) traffic situation

The questionnaire also contained an abbreviated Dutch version (Verschuur, 2003) of the Driver Behaviour Questionnaire (DBQ, see Parker, Reason, Manstead, & Stradling, 1995), which relates to the strategic level, as well as the manoeuvring level. The 8 items of the DBQ indicating violating behaviour were used to measure the effect of inadequate self-assessment on violating behaviour.

On-road driving assessment. A driving assessment was conducted, consisting of half an hour driving on different types of road. Examiners rated drivers on their ability to drive safely on a scale from 0 to 10, 5.5 being the pass-fail criterion in a real driving test. As in a regular driving test, this general mark for "safe driving" is a combination of many different skills on different levels of the driving task (e.g., to *manoeuvre* safely through traffic, a driver must possess a certain level of vehicle *control*).

Data analysis

Self-assessment was computed by comparing the examiner's rating for safe driving with the participants' confidence-rating. Chi-square analysis was used to test for significant differences in frequencies. Analysis of variance was used to test for significant effects (α = .05) on the DBQ items and the adaptation score. Besides significance of the results, the effect size (Partial èta squared, η^2) is also reported, with $\eta^2 \approx .01$ as a small, $\eta^2 \approx .06$ as a medium, and $\eta^2 \approx .14$ as a large effect size (Cohen, 1988).

RESULTS

The variable "driver confidence" was used to distinguish two driver groups: a) high-confidence drivers (drivers who said they were (very) confident); and b) low-confidence drivers (drivers whose responses were neutral or not (at all) confident). Next, the examiner's "safe driving" score was used to divide the drivers into two groups: a pass group, consisting of those who would have passed the driving test (scoring 5.5 or higher) and a fail group consisting of those who would have failed (scoring less than 5.5). Table 1 shows a cross tabulation of "driver confidence" and the examiner's "safe driving" score. This results in three groups: a) 88 (82 + 6) well-calibrated drivers (share the same opinion about their driving skills), b) 20 insecure drivers (are not confident about their driving skills, but passed the driving assessment) and c) 22 overconfident drivers (are confident about their driving skills, but failed the driving assessment). The reported results in this paper are fairly robust for the pass-fail criterion for "safe driving" (that is, the results are similar when a pass-fail criterion of, for example, 6 is used).

Table 1. Calibration groups

	Safe driving		Total
Confidence	Passed	Failed	
Confident	82[a]	22[c]	104
Insecure	20[b]	6[a]	26
Total	102	28	130

Table 2 shows some characteristics of the three calibration groups. Chi-square analysis indicated that the group of experienced drivers consisted of significantly more well-calibrated drivers (85%) than the group of novice drivers (58%); ($\chi^2_{(1, N = 130)} = 10.21$; $p < .01$). The difference between males and females in the three calibration groups was also significant ($X^2_{(2, N = 130)} =$ 8.83; $p < .05$); 77% of the males were well-calibrated drivers; of the female drivers only 57% were well calibrated. Female drivers were significantly overrepresented in the insecure group (25% of the female drivers compared to 7% of the male drivers; $\chi^2_{(1, N = 130)} = 7.91$; $p < .01$). There was no gender effect in the overconfident group.

Table 2. Characteristics of calibration groups

	n	Well calibrated	Insecure	Overconfident
Group				
Experienced	47	85%	11%	4%
Novice	83	58%	18%	24%
Gender				
Male	70	77%	7%	16%
Female	60	57%	25%	18%

Adaptation

Strategic level. The results show that overall, 25% of the drivers reported "not driving because of adverse conditions"; no difference was found between the well-calibrated group and the other two groups. A larger proportion of drivers (44%) reported having avoided a difficult intersection. However, further analyses showed that drivers avoided such intersections because of traffic congestion not because of traffic complexity. Consequently, there was no difference between calibration groups. There was also no individual effect of "driver confidence" or performance on the driving assessment ("safe driving"), on these variables.

Manoeuvring level. ANOVA shows that there was a moderate ($\eta^2 = .09$) and significant difference between calibration groups ($F_{2,124} = 6.15$; $p < .01$) in the responses on the Adaptation Test. Post Hoc Bonferroni tests indicated that overconfident drivers adapted their speed to the complex situation significantly less (26% correct responses; $p < .01$) than well-calibrated drivers (42% correct responses) and insecure drivers (36% correct responses). No significant interaction effect was found between calibration groups and gender.

Safe drivers, as determined in the driving assessment, reported reduced speed more often in the complex situations (41% correct responses) compared to unsafe drivers (31% correct responses; $F_{1,126} = 4.56$; $p < .05$). No effect of "driver confidence" was found on the Adaptation Test.

Violations. The 8 items of the DBQ indicating violating behaviour (Table 3) were used to see whether well-calibrated drivers reported less dangerous decisions. MANOVA indicated that there was a large ($\eta^2 = .13$) and significant difference between calibration groups ($F_{16,236} = 2.11$; $p < .01$) in their reported violating behaviour. Tests of between-subject effects indicate that, except for being angered by another driver (item 4), and having an aversion to a class of road users (item 7), all items differentiated significantly between calibration groups. The overconfident group generally reported more violating behaviour than the well-calibrated and insecure group. Table 3 shows the mean score for each group on the 8 DBQ items indicating how often they would indulge in that behaviour (0 = Never; 5 = Nearly all the time). Significant deviations ($\alpha = .05$) from the well-calibrated group, based on Bonferroni Post Hoc tests, are indicated by an asterisk.

Table 3 Mean score on DBQ items

		Mean score		
		Well calibrated	Insecure	Overconfident
1.	Become impatient with a slow driver in the outer lane and overtake on the inside	2.0	2.1	2.8*
2.	Drive especially close to the car in front as a signal to its driver to go faster or get out of the way	1.9	2.0	2.6*
3.	Cross a junction knowing that the traffic lights have already turned against you	1.7	1.8	2.3*
4.	Angered by another driver's behaviour, you give chase with the intention of giving him/her a piece of your mind	1.2	1.1	1.1
5.	Disregard the speed limits late at night or early in the morning	2.8	3.6	3.5
6.	Drive even though you realise that you may be over the Legal blood-alcohol limit	1.2	1.1	1.6*
7.	Have an aversion to a particular class of road user, and indicate your hostility by whatever means you can	1.5	1.6	2.0
8.	Get involved in unofficial 'races' with other drivers	1.4	1.2	1.9

Note: * indicates a significant difference ($\alpha = .05$) with well-calibrated drivers

No main effect of gender was found on the DBQ-items. However, the interaction of gender and calibration group was significant ($F_{16,236} = 2.16$; $p < .01$). Tests of between–subject effects showed that the interaction effect was exclusively caused by ‘driving over the Legal blood-alcohol limit’ (item 6). Only in the overconfident group, male drivers reported significantly more drink and drive behaviour than females. In the well-calibrated and insecure group, no effect of gender was found.

The items of the DBQ were additionally analysed for the effects of “driver confidence” and performance on the driving assessment (“safe driving”). Confident drivers reported more violations than less confident drivers ($F_{8,119} = 2.47$; $p < .05$), but no effect of “safe driving” was found.

DISCUSSION

As mentioned in the introduction, research has suggested that, compared to experienced drivers, novice drivers are poor at calibration and that inadequate self-assessment, and specifically, overestimation of skills, results in inadequate adaptation of task demands. To measure self-assessment of skills, drivers’ self-reported confidence was compared with actual driving performance. The cross tabulations resulted in three groups: a) well-calibrated drivers, b) overconfident drivers and c) insecure drivers. Only half of the novice drivers were “well calibrated” as opposed to 85% of the experienced drivers. As expected, a relatively large proportion of novices (24%) belong to the “overconfident” group. This proportion is particularly large, when taking into account the modest amount of driving experience in the novice driver group.

The main hypothesis of this paper, that inadequate self-assessment of skills may lead to dangerous behaviour (irrespective of experience level), was supported by the finding that overconfident drivers generally reported more violating behaviour. Our results also show that overconfident drivers reported less instances of adaptation of driving speed to the complexity of traffic situations as measured with the Adaptation Test. Although insecure drivers performed somewhat worse on the Adaptation Test, the percentage of correct responses was not significantly different from the well-calibrated drivers. There was no evidence found to suggest that inadequate self-assessment of skills affects adaptation on the strategic level of the driving task. This could have been a result of the choice of questions to measure adaptation on this level. There were only two questions, of which one seemed to have measured something other than it was intended to measure (i.e., drivers considered congestion as a reason to avoid a difficult intersection, instead of complexity of the traffic situation).

The construction of the calibration groups was based on a comparison of self-reported confidence (“driver confidence”) and the examiners “safe driving” scores. It may be argued that reported differences between calibration groups are caused solely by, for example, “safe driving,” and that “driver confidence” did not contribute to the reported differences. It is difficult to eliminate this alternative explanation, because the number of participants is too few to have a reasonable power for tests of interaction between confidence and examiners’ opinion. As a result, no significant interaction effects were found. However, we did not find any evidence that either “driver confidence” or “safe driving” could explain all the results. For example, the violation items of the DBQ indicated a significant main effect for "driver confidence" but there was no significant main effect for "safe driving." In contrast, for the Adaptation Test, a significant effect for "safe driving" was found but not for "driver confidence." These results suggest that the combination of "safe driving" and "driver confidence" (i.e., self-assessment of skills) explains more variation than both factors separately.

Prior research has indicated calibration as a relevant factor for safe driving, and has linked calibration to the high crash rate of young, novice drivers. The current study provides evidence for the behavioural consequences of "overconfidence." Whether, how, and when calibration improves in the first months of licensing is subject of further study.

REFERENCES

Brown, I.D. & Groeger, J.A. (1988). Risk perception and decision taking during the transition between novice and experienced driver status. *Ergonomics, 31*(4), 585-597.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale NJ: Lawrence Erlbaum Associates.

De Craen, S., Twisk, D.A., Hagenzieker, M.P., Elffers, H.E., & Brookhuis, K.A. (2007). *The development of a method to measure speed adaptation to traffic complexity: Identifying novice, unsafe and overconfident drivers*. Manuscript submitted for publication.

Fuller, R. & Santos, J.A. (2002). *Human Factors for Highway Engineers*. Oxford: Elsevier Science Ltd.

Fuller, R. (2005). Towards a general theory of driver behaviour. *Accident Analysis and Prevention, 27*(3), 461-472.

Gregersen, N.P. (1995). *Prevention of road accidents among young novice car drivers.* Doctoral dissertation, Linköping University, Linköping, Sweden.

Kuiken, M.J. & Twisk, D.A.M. (2001). *Safe driving and the training of calibration* (Report no. R-2001-29). Leidschendam, the Netherlands: SWOV Institute for Road Safety Research.

Mathews, M.L. & Moran, A.R. (1986). Age differences in male drivers' perception of accident risk: The role of perceived driving ability. *Accident Analysis & Prevention, 18*, 299-313.

Michon, J.A. (1984). A critical view of driver behavior models: What do we know, what should we do? *Human behavior and traffic safety: Proceedings of a general motor symposium on human behavior and traffic safety*, Michigan, 516-520.

Parker, D., Reason, J.T., Manstead, A.S.R., & Stradling, S.G. (1995). Driving errors, driving violations and accident involvement. *Ergonomics, 38*(5), 1036-1048.

Rasmussen, J. (1986). *Information processing and human-machine interaction; an approach to cognitive engineering*. Amsterdam, the Netherlands: North-Holland.

Reason, J. (1990). *Human Error*. Cambridge: Cambridge University Press.

Verschuur, W.L.G. (2003). *Onderzoek uitgevoerd met de 'Driver Behaviour Questionnaire'; eindrapportage* (Werkgroep Veiligheid/R-03/68; in Dutch). Leiden, The Netherlands: Rijksuniversiteit Leiden.