How indicative is a self-reported driving behaviour profile of police registered traffic law offences?

Martinussen, Laila Marianne; Møller, Mette; Prato, Carlo Giacomo; Haustein, Sonja

Published in: Accident Analysis & Prevention

Link to article, DOI: 10.1016/j.aap.2016.10.031

Publication date: 2017

Document Version
Peer reviewed version

How indicative is a self-reported driving behaviour profile of police registered traffic law offences?

L. M. Martinussen a, M. Møller a, C. G. Prato b, & S. Haustein a*

a Technical University of Denmark, Department of Management Engineering, DK-2800 Kgs. Lyngby, Denmark
b School of Civil Engineering, The University of Queensland, St. Lucia 4072, Brisbane, Australia

*Corresponding author. Tel. +45254727
E-mail address: sonh@dtu.dk
Abstract

Although most motorised countries have experienced massive improvements in road safety over the last decades, human behaviour and differences in accident risk across sub-groups of drivers remains a key issue in the area of road safety. The identification of risk groups requires the identification of reliable predictors of safe or unsafe driving behaviour. Given this background, the aim of this study was to test whether driver sub-groups identified based on self-reported driving behaviour and skill differed in registered traffic law offences and accidents, and whether group membership was predictive of having traffic law offences. Sub-groups of drivers were identified based on the Driver Behaviour Questionnaire (DBQ) and the Driver Skill Inventory (DSI), while traffic offences and accidents were register-based (Statistics Denmark). The participants (N=3683) were aged 18-84 years and randomly selected from the Danish Driving License Register. Results show that the driver sub-groups differed significantly in registered traffic offences but not in registered accidents. In a logistic regression analysis, the sub-group “Violating unsafe drivers” was found predictive of having a traffic offence, even when socio-demographic variables and exposure were controlled for. The most important predictive factor, however, was having a criminal record for non-traffic offences, while gender, living without a partner, and being self-employed also had a significant effect. The study confirms the use of the DBQ and DSI as suitable instruments for predicting traffic offences while also confirming previous results on accumulation of problematic behaviours across life contexts. The finding that driver sub-groups did not differ in registered accidents supports the recent research activities in finding and modelling surrogate safety measures.

Keywords: Segmentation, Traffic law offences, Accidents, Criminal record, DBQ, DSI
1. Introduction

Human behaviour is a key factor in 80-90% of road traffic accidents (e.g., Rothengatter, 1997; Shinar, 2007). For the development of effective preventive measures, it is therefore crucial to know which types of driving behaviours are problematic in the context of road safety and which sub-groups of drivers perform these behaviours.

Several studies have identified sub-groups of drivers using self-report measures. In a recent study, Martinussen et al. (2014) applied two self-report measures to identify sub-groups of drivers that differ in their propensity to drive in aberrant ways: the Driver Behaviour Questionnaire (DBQ, Reason et al., 1990) and the Driver Skill Inventory (DSI, Lajunen & Summala, 1995). The study identified four driver sub-groups of which two stood out as potentially more unsafe than the other two sub-groups: the “Violating unsafe drivers” and the “Unskilled unsafe drivers”. These two groups reported the highest levels of aberrant driving behaviour, and lowest technical driving skills or safety skills, or both. They also reported significantly more accidents and fines. As comparably safe driver groups “Skilled safe drivers” and “Low confidence safe drivers” were identified (for details, see Martinussen et al., 2014).

However, this study did not answer the question whether the group differences based on self-reported data were also related to traffic offences and accidents as reported by the police. This question is relevant, as self-reports on driving behaviour and accident involvement have been criticised as a method because persons may modify their answers for social desirability reasons, may remember episodes incorrectly (memory bias), and may want to report consistently across related measures (common method variance, CMV) (af Wåhlberg, 2010; af Wåhlberg et al, 2011). More specifically, the usefulness of the DBQ has been questioned because of its limited ability to predict accidents (af
Wåhlberg et al, 2011; af Wåhlberg & de Winter, 2012). In a recent paper, af Wåhlberg et al. (2015) concluded that DBQ’s predictability of accidents was driven by an exposure effect: drivers with a high number of violations did not violate more, they just drove more while violations per kilometre were not higher, which stresses the necessity to control for mileage when comparing self-reported driver behaviour. af Wåhlberg et al. (2015) suggested further research was needed where DBQ data should be compared with registered data, thereby not susceptible to CMV.

With the unique opportunity in Denmark of combining register data from Statistics Denmark to survey data such as the DBQ and DSI on a representative sample of the population, this study examined whether the differences between driver sub-groups as identified by Martinussen et al. (2014) were observed also when comparing police registered traffic offences and accidents. Moreover, we examined to what extent possible differences between the four sub-groups of drivers in registered traffic offences could be explained by differences in their socio-demographic characteristics (i.e., age, gender, living with a partner, income, education, living in Copenhagen, and car ownership as well as having a criminal record for non-traffic offences) and mileage; that means whether group membership was (still) predictive of traffic violations, when demographics and exposure were controlled for. More specifically, we formulated the following hypotheses:

**Hypothesis 1**: The driver sub-groups identified as unsafe based on self-report data (‘Violating unsafe drivers’; ‘Unskilled unsafe drivers’) have more registered traffic law offences and accidents than the two safe groups.
Hypothesis 2: When predicting traffic law offences based on group membership, belonging to one of the unsafe groups has still a significant effect on registered traffic law offences, when socio-demographic factors are controlled for.

Hypothesis 3: When controlling for exposure, the effect of “Violating unsafe drivers” (the group with the highest mileage) is no longer significant.

The results were expected to shed light on the validity of the identified driver sub-groups and thereby also indirectly on the instruments the groups were based upon, namely the DBQ and DSI. In addition, the analyses were expected to reveal which socio-demographic characteristics were predictive for registered traffic law offences, providing additional knowledge for the design and targeting of preventive measures.

2. Method

2.1. Participants

The sample consists of 3683 persons who took part in a survey on driver behaviour and could afterwards be matched with data from Statistics Denmark. Originally, 11,004 individuals aged 18-84, randomly drawn from the Danish Driving License Register (stratified by age and gender) received a letter announcing the study together with the questionnaire, a freepost return envelope, and a web address to return the questionnaire online if preferred. Two reminders were sent. The response rate was 44 percent. Of the 4849 respondents who returned a questionnaire, 941 (19%) had to be excluded as they did not complete the full questionnaire and of these 225 (5%) had to be excluded as they could not be matched with data from Statistics Denmark, resulting in the final sample of 3683. Additional details
about the sampling process can be found in Martinussen et al. (2013; 2014), while characteristics of the sample can be found in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Categories</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>18-24 years old</td>
<td>10.1%</td>
</tr>
<tr>
<td></td>
<td>25-34 years old</td>
<td>11.9%</td>
</tr>
<tr>
<td></td>
<td>35-44 years old</td>
<td>15.5%</td>
</tr>
<tr>
<td></td>
<td>45-54 years old</td>
<td>17.3%</td>
</tr>
<tr>
<td></td>
<td>55-64 years old</td>
<td>17.5%</td>
</tr>
<tr>
<td></td>
<td>65-74 years old</td>
<td>16.1%</td>
</tr>
<tr>
<td></td>
<td>75-84 years old</td>
<td>11.6%</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>47.6%</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>52.4%</td>
</tr>
<tr>
<td>Household</td>
<td>Living alone</td>
<td>21.9%</td>
</tr>
<tr>
<td></td>
<td>Living in a multi-person household</td>
<td>78.1%</td>
</tr>
<tr>
<td>Living in Copenhagen</td>
<td>Yes</td>
<td>93.3%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>6.7%</td>
</tr>
<tr>
<td>Education</td>
<td>Low</td>
<td>65.1%</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>23.1%</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>8.2%</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>3.6%</td>
</tr>
<tr>
<td>Employment status</td>
<td>Employee</td>
<td>59.3%</td>
</tr>
<tr>
<td></td>
<td>Self-employed</td>
<td>5.4%</td>
</tr>
<tr>
<td></td>
<td>Retiree</td>
<td>28.5%</td>
</tr>
<tr>
<td></td>
<td>Unemployed</td>
<td>4.8%</td>
</tr>
<tr>
<td></td>
<td>In education</td>
<td>2.0%</td>
</tr>
<tr>
<td>Car ownership</td>
<td>Yes</td>
<td>67.6%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>32.4%</td>
</tr>
<tr>
<td>Mileage (self-reported)</td>
<td>until 6000 km/year</td>
<td>29.4%</td>
</tr>
<tr>
<td></td>
<td>6000-12000 km/year</td>
<td>25.5%</td>
</tr>
<tr>
<td></td>
<td>12000-18000 km/year</td>
<td>15.8%</td>
</tr>
<tr>
<td></td>
<td>18000-24000 km/year</td>
<td>10.8%</td>
</tr>
<tr>
<td></td>
<td>more than 24000 km/year</td>
<td>18.5%</td>
</tr>
<tr>
<td>Traffic offences</td>
<td>Yes</td>
<td>10.8%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>89.2%</td>
</tr>
<tr>
<td>Criminal record</td>
<td>Yes</td>
<td>2.7%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>97.3%</td>
</tr>
</tbody>
</table>
2.2. Measures

Sub-groups of drivers were identified based on a cluster analysis of self-reported answers to DBQ and DSI. The DBQ was used to assess aberrant driver behaviour by asking how often the drivers performed violations, errors and lapses on a six-point scale (0 = never, 5 = nearly all the time) across different driving situations (for details see Martinussen et al., 2014; Reason et al., 1990).

The DSI was used to assess perceptual-motor skills and safety skills by asking drivers to assess how skilful they considered themselves to be compared with the average driver across different driving situations. A five-point scale (0 = well below average, 4 = well above average) was used (for details see Lajunen & Summala, 1995; Martinussen et al., 2014). Based on their answers to the DBQ and the DSI, the participants were clustered into four groups of drivers (“Skilled safe drivers”, “Violating unsafe drivers”, “Unskilled unsafe drivers”, “Low confidence safe drivers”) as described in the introduction and in more detail in Martinussen et al. (2014). The names of the clusters reflect their scores on the two scales (e.g., skilled safe drivers = high score on skills/DSI and low score on aberrant driving behaviour/DBQ).

In this study, for each participant register based information was derived from Statistics Denmark and added to the survey data of the respective person. The information included demographic information (income, education, family status, and car ownership), accident involvement (police registered injury and fatal accidents), registered traffic law offences, and having a criminal record resulting from non-traffic offences. The register based information on demographics was taken from the year in which the participant took part in the survey (2010 or 2011). Mileage was included in the survey and is thus self-reported. For accident involvement, traffic offences, and criminal record, three
dummy variables were created which indicated whether a participant did or did not have one or more
incidents of each category in the period 2007-2012.

2.3. Statistical analysis
Chi-square tests were used to test for significant differences between the driver sub-groups with regard
to police registered accident involvement and traffic offences.

Three logistic regression analyses were conducted to predict the likelihood of having a police
registered traffic offence. The three analyses added on the predictors in order to test the hypotheses.
Specifically, the first model included as predictors the belonging to the sub-group of drivers: “Violating
unsafe drivers”, “Unskilled unsafe drivers”, or “Low confidence safe drivers” (Model 1). In the second
model, age, gender, living with a partner, income, education, living in Copenhagen, car ownership, and
having a criminal record for non-traffic offences were added (Model 2), while the final model
controlled additionally for car mileage (Model 3).

3. Results
3.1. Sub-groups’ traffic offences and accident involvement
Table 2 shows the percentages of drivers with a police registered traffic offence within each sub-group
of drivers. According to the results, more than three times as many “Violating unsafe drivers” were
registered for traffic offences than “Low confidence safe drivers”, and more than two times as many as
persons in the other two sub-groups. Other than expected, “Unsafe unskilled drivers” did not emerge
with a higher number of traffic offences. In addition, Table 2 shows that there was no significant
difference in accident involvement: in each sub-group approximately 1% were involved in a traffic
accident involving injuries in the period 2007-2012. Thus, Hypothesis 1 has to be partly accepted for what concerns the traffic offences, and partly rejected for what concerns the accident involvement.

Table 2: Percentages of drivers with a police registered traffic offence and accident involvement within sub-groups of drivers

<table>
<thead>
<tr>
<th>Sub-groups of drivers</th>
<th>Skilled safe drivers</th>
<th>Violating unsafe drivers</th>
<th>Unskilled unsafe drivers</th>
<th>Low confidence safe drivers</th>
<th>Total</th>
<th>Chi² test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recorded traffic law offences</td>
<td>9.3%</td>
<td>22.4%</td>
<td>8.3%</td>
<td>7.0%</td>
<td>10.8%</td>
<td>Chi² (df=3, n=3787) =118.76, p &lt; .001</td>
</tr>
<tr>
<td>Recorded accidents</td>
<td>1.1%</td>
<td>1.0%</td>
<td>1.2%</td>
<td>1.2%</td>
<td>1.1%</td>
<td>Chi² (df=3, n=3787) =0.10, p &gt; .10</td>
</tr>
</tbody>
</table>

Note: df = degrees of freedom; n = number of observations; p = significance level of the parameter estimate

3.2. Factors related to a traffic law offence

So far, we have shown that “Violating unsafe drivers” are more likely to have a police registered traffic offence as compared to the other three driver groups. However, it remains unclear if the group differences are due to differences in self-reported driving behaviour/skill or due to the composition of the groups in terms of other variables. Hence, we estimated three logistic regression models: in Model 1, only the driver groups were included; in Model 2, additional demographic variables were controlled for, as well as having a non-traffic criminal record; in Model 3, the mileage was also controlled for.

The likelihood ratio test (LRT) statistic for Model 2 when compared to Model 1 revealed that there was a significant improvement in the model performance ($LRT = 123.75, df = 12, p < 0.001$), and similarly the same statistic for Model 3 when compared to Model 2 showed another significant improvement in the performance ($LRT = 36.77, df = 1, p < 0.001$). Table 3 shows the odds ratios and their significance level for the factors that, in the year of the survey (2010/2011), were significantly related to having a registered traffic offence in the analysed period 2007-2012.
In line with the descriptive analysis, belonging to the group of “Violating unsafe drivers” was significantly related to having a registered traffic offence, while this was not the case for the membership to one of the other driver sub-groups. When additional factors were included in the analyses (Model 2), the group membership was still significant confirming Hypothesis 2 with regard to the “Violating unsafe drivers”. Yet, in Model 2, having a criminal record for other offences became the most important factor: the likelihood of getting a traffic offence within the considered 6-years period was almost three times higher for persons with a non-traffic related criminal record than for people with no offence. Belonging to the sub-group of “Violating unsafe drivers” was the second most important factor. In addition, several socio-demographic variables were significantly related to having a traffic offence: living alone increased the likelihood by 58%, being self-employed by 67%, and being a car-owner by 40%. In contrast, being female reduced the probability by 35%. The finding that both age and age squared became significant shows that there is a non-linear relationship between age and having a police registered traffic offence with a general increase by almost 5% for every year of age smoothed downwards by the quadratic effect. However, when mileage was included in the analysis (Model 3), age was no longer significant. All other variables that were significant in Model 2 remained significant when controlling for mileage, so Hypothesis 3 can be rejected.

Table 3: Logistic regressions modelling the likelihood of getting registered for a traffic offence within the period 2007-2012.

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>O.R.  p</td>
<td>O.R.  p</td>
<td>O.R.  p</td>
</tr>
<tr>
<td>Violating unsafe drivers</td>
<td>2.789 0.000</td>
<td>1.846 0.000</td>
<td>1.718 0.000</td>
</tr>
<tr>
<td>Unskilled unsafe drivers</td>
<td>0.875 0.413</td>
<td>0.908 0.567</td>
<td>0.913 0.594</td>
</tr>
<tr>
<td>Low confidence safe drivers</td>
<td>0.732 0.041</td>
<td>0.853 0.312</td>
<td>0.895 0.486</td>
</tr>
<tr>
<td>Age</td>
<td>1.046 0.041</td>
<td>1.025 0.260</td>
<td></td>
</tr>
<tr>
<td>Age²</td>
<td>0.999 0.009</td>
<td>1.000 0.081</td>
<td></td>
</tr>
<tr>
<td>Gender (female)</td>
<td>0.654 0.001</td>
<td>0.741 0.018</td>
<td></td>
</tr>
</tbody>
</table>
### 4. Discussion

The purpose of the present study was to test whether driver sub-groups differing in self-reported driving behaviour and skill also differed in registered traffic law offences and accidents, and whether sub-group membership predicted having police-registered traffic offence, even when socio-demographic variables and exposure were controlled for.

As the descriptive statistics showed, the percentage of persons who had a registered traffic offence was higher within the group of “Violating unsafe drivers”, while there were no distinct differences between the other three driver sub-groups. This result was confirmed by the results of the regression analyses. Contrary to the conclusions of af Wåhlberg et al. (2015), the effect of the “Violating unsafe drivers” in the regression analyses remained significant when controlling for exposure. This finding highlights the “Violating unsafe drivers” as the most unsafe driving sub-group.
and supports the predictive value of the DBQ and the DSI and is further supported by naturalistic
driving studies, which show that self-reports on driving behaviour correlate to risky driving behaviour.
For example, speeding with sudden unidirectional acceleration was observed for drivers with high
DBQ factor scores (Zhao et al., 2012), speeding in daylight conditions was associated with high DBQ
violations (Helman & Reed, 2015), and risky driving behaviour was correlated to high self-reported
reckless and angry driving styles (Taubman – Ben-Ari et al., 2016).

An unforeseen result was, however, that being part of “Unskilled unsafe drivers” was not
predictive of having traffic law offences. However, this might be due to the fact the “Unskilled unsafe
drivers” behaviour is characterised by unintentional errors and lapses, which to a lesser extent lead to
registered violations of the traffic law compared to the “Violating unsafe drivers” behaviour that is
characterised by intentional violations (which per definition includes behaviours such as speeding and
drunk driving). One might also conclude the combination of the DBQ/DSI is sensitive enough to really
separate between intentional violations (such as speeding and drunk driving) and unintentional errors
and lapses, which to a lesser extent lead to a traffic offence. The results indicate that preventive efforts
should primarily focus on the “Violating unsafe drivers”.

Contrary to expected, the results showed that all groups had the same amount of registered
accidents. Thus, the DBQ/DSI combination does not seem to predict recorded accidents well, which
also indicates that the two unsafe sub-groups identified by Martinussen et al. (2014) may not be the
ones having more accidents. It has been argued that the DBQ is of little worth if it does not predict
accidents (af Wåhlberg et al., 2011). However, a difference should be noted between committing a
traffic violation and being involved in an accident: while the former clearly implies that the driver
engaged in a behaviour that resulted in a violation, the latter does not. Arguably, accident involvement
might result from the wrongful behaviour of other drivers. Thus, it is possible, that although the amount of registered accidents did not differ between the groups, the type of accidents differed with regard to whom or what caused the accident. In a study on young moped rider accidents (Møller & Haustein, 2016), it has been shown that in 27% of the accidents the behaviour of the other party involved caused the accident. Further, it is well known that in most cases a coincidence of other (non-behavioural) factors, which are outside the drivers’ control, play a significant role (Elvik, 2010). The data included in this study did not allow detailed accident analysis, but as we agree with af Wåhlberg et al. (2011) that instruments, such as the DBQ, should be predictive of unsafe driving behaviour, and that this should mirror some kind of registered safety/risk measure/data, further studies exploring possible difference in accident type and behavioural influence between the four groups are recommended to see if the combination of DBQ and DSI predicts accidents at a more detailed level. Nevertheless, we consider it an important finding that DBQ and DSI are predictive for traffic offences.

With regard to the development of targeted preventive measures, the results of this study confirm that targeted safety measures preferably should take multiple preventive factors into account. Thus, the results of the regression analyses confirmed the relevance of previously identified factors such as age, gender, living without a partner (e.g., Møller et al., 2015) and having a criminal record for other offences (e.g., LaBrie et al., 2007). In this study, being female reduced the probability of having a traffic offence by 35%, though this was reduced to 26% when taking women’s lower mileage into account. This is in line with results from previous studies showing that being male was a predictor of driving violations (e.g., Özkan & Lajunen, 2005), and of having a greater risk of being involved in an accident than females (e.g., Hansen & Jensen, 2012). The fact that living without a partner stood out as a predictive value of traffic offences also confirms previous results showing that family status is
indicative of involvement in problem behaviours, for instance living without a partner have been found to be predictive of being a drunk driving recidivist (Møller et al., 2015). Having a criminal record for non-traffic offences stood out as the most predictive factor. This confirms that there is a relationship between peoples’ lifestyle in general and their driving behaviour (e.g., Roach, 2007; Møller & Sigurdardottir, 2009; Møller et al., 2015), leading problem behaviour such as violations to accumulate, a phenomenon which has previously been denoted the problem behaviour syndrome (e.g., Shope et al., 2003). This knowledge is of key importance for the development of targeted preventive measures and implies that measures that only focus on traffic violation as the problem behaviour may be short-sighted as traffic offences can rather be seen as part or consequence of broader problem behaviour of a specific population sub-group.

Finally, it should be noted that participation in the postal survey was voluntary. Therefore, possible bias due to self-selection among the respondents is possible. Non-response can be cognitively based (e.g. insufficient reading/writing skills) as well as motivationally based (e.g. lack of interest in the subject) (Beatty & Herrmann, 2002). In relation to this study it is possible that motivationally based self-selection bias has influenced the results. Thus, it is possible that persons belonging to the group of “Violating unsafe drivers” were less motivated to participate in a study on road safety, whereas having been involved in an accident may have motivated participation among members from the other groups, thereby blurring the group differences with regard to accident involvement. Unfortunately, such non-response analysis was not possible, but it may have biased the results regarding DBQ and DSI as predictors of accident involvement. In addition, self-report data can be biased by factors such as social desirability (Lajunen & Summala, 2003), response, memory and/or hindsight bias (Roese & Vohs,
However, the sample is quite large and the amount of traffic offences in the register suggests that not only safe drivers participated in the survey as could have been hypothesized.

5. Conclusion

The characteristics of “Violating unsafe drivers” found by Martinussen et al. (2014) were also confirmed by register traffic offence data even when socio-demographic factors and exposure were controlled for, thereby indicating real behavioural differences between this group and the three others. However, this was only confirmed with traffic offences and not accidents. Due to the multi-causal nature of accidents, we question accidents as the only and main indicator of dangerous driving. Because accidents are rare in western countries, surrogate measures of dangerous driving, such as traffic offences might be a more reliable predictors of safe and unsafe driving behaviour. The present results indicate that DBQ and DSI are reliable predictors of traffic offences, which appears to be in line with the core aim of the scales.

Preventive efforts should primarily focus on “Violating unsafe drivers”, where future studies should look into the most salient behaviours in order to target suitable preventive measures. However, the result that the strongest predictor for having a traffic offence was having a non-traffic offence indicates that behaviour is consistent across situations and leading problem behaviour such as violations to accumulate, in particular among male singles. In terms of prevention, the results of this study therefore indicate that measures should not focus solely on punishment of single traffic violations, but could also profitably address issues related to crime and social marginalisation.
Acknowledgements

Data management support of Allan Hansen is greatly appreciated. We further acknowledge the contribution of the reviewers to sharpen an initial version of the paper.

This work was supported by Innovation Fund Denmark.


Helman, S., & Reed, N. (2015). Validation of the driver behaviour questionnaire using behavioural data from an instrumented vehicle and high-fidelity driving simulator. Accident Analysis & Prevention, 75, 245-251.


