Validation of in-car observations, a method for driver assessment

Magnus Hjälmdahl *, András Várhelyi

Department of Technology and Society, Lund University, Lund, Sweden

Received 18 February 2003; received in revised form 18 September 2003; accepted 22 September 2003

Abstract

An in-car observation method with human observers in the car was studied to establish whether observers could be trained to observe safety variables and register driver’s behaviour in a correct and coherent way, and whether the drivers drove in their normal driving style, despite the presence of the observers. The study further discussed the observed variables from a safety perspective. First three observers were trained in the observation method and on-road observations were carried out. Their observations were then compared with a key representing a correct observation. After practising the observation method the observers showed a high correlation with the key. To establish whether the test drivers drove in a normal way during the in-car observations, comparisons of 238 spot-speed measurements were carried out. Driver’s speeds when driving their own private cars were compared with their speeds during the in-car observations. The analysis showed that the drivers drove in the same way when being observed as they did normally. Most of the variables studied in the in-car observations had a well documented relevance to traffic safety. Overall, in-car observation was shown to be a reliable and valid method to observe driver behaviour, and observed changes provide relevant data on traffic safety.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Driver behaviour; In-car observations; Safety assessment; Method validation

1. Introduction

A variety of observational methods are currently being used in order to investigate the possibility that the use of advanced driver assistance systems (ADAS) may lead to negative behavioural modifications, see e.g. the negative effect of ABS-brakes indicated by behavioural observations of
drivers and by accident statistics (Aschenbrenner et al., 1993; Highway Loss Data Institute, 1994). Typically, driver behaviour is studied using driving simulators, instrumented vehicles or human observers inside the vehicle. While all these methods have their advantages, their disadvantages are that they can generate data that is not always reliable or valid.

Simulators and instrumented vehicles are useful for studying static driver behaviour; for instance speed, lane keeping and eye movements. They can also simulate critical conditions and even push the test-drivers so hard that they crash. The same conditions can be simulated over and over and testers can thus select specific situations and study them thoroughly. The disadvantage of simulators is that their use makes it hard to create situations in which drivers feel as though they are driving in real traffic. The instrumented vehicle has the advantage that it can be used in real traffic and, if the instruments are hidden and unobtrusive, the drivers will get very close to their normal behaviour. However, with this method it is difficult to get a correct view of interactions and communication with other road users, possible violations and conflicts.

One promising way to achieve an understanding of how and why a situation occurs is to have a human observer in the vehicle. For this purpose the in-car observation method was developed by Risser (1985). Originally the method was meant to be used to appraise learning drivers, but it is also suitable for research on driver behaviour in real traffic. The original method used two observers in the car, studying the driver along a specified test route; in this study we have used an instrumented vehicle in addition to the two observers. One of the observers registers standardised variables such as speed adaptation at junctions/obstacles, lane change, interaction with other road users, etc. While in the original observation method, this observer also registered speed, lateral position and distance to vehicle in front, in this study we have used the equipment in the instrumented vehicle instead. The other observer makes free observations, about, for example conflicts, communication, interaction and special events. It is important to determine whether the observer has any effect on the driver, thereby making it impossible to say whether the observed behaviour is a result of the ADAS in question or an observer effect.

A few studies have dealt with this issue of observer effect and there are some differences in the results. Höfner (1967) found that the behaviour of moped riders did not change when they knew that they were being observed. On the other hand, Rathmayer et al. (1999) found that subjects, driving an instrumented car with an experiment leader, had a 1–2 kph lower mean speed when the experiment leader was present. They further found that acceleration and deceleration were smoothed down and lateral acceleration was reduced. Quimby (1988) found that if the drivers were given a subsidiary task to perform, which they believed to be the purpose of the experiment, they did not suspect their driving skill was being studied and therefore they did not change their driving behaviour.

In order to determine the effect of the observers on the drivers, we compare the drivers’ behaviour when they drive their own cars with no observers and no instrumentation (except for an unobtrusive data-logger), to when they drive an instrumented car with two observers in the car. Such a study was made possible by the large-scale trial with Intelligent Speed Adaptation (ISA) in Lund (Várhelyi et al., 2002).

The ISA system studied in the trial was the so-called active accelerator pedal. The system is designed to help the driver to keep within the legal speed-limit by tactile feedback in the accelerator which tells the driver when the speed-limit is reached. In total, there were 284 vehicles in Lund equipped with the active accelerator and a data logger to register driving data continuously.
The test drivers drove for the first month, without the ISA system activated, to generate before data, which is regarded as the drivers’ normal driving behaviour.

This study aims to investigate how valid and reliable the in-car observation method is in observing driver behaviour, and to establish the relation between the studied variables and traffic safety. It further aims to establish a standardised procedure for training and validating of the observers.

2. Method

2.1. Test driver selection

The total population of 284 ISA-drivers in Lund were randomly selected from the motor vehicle registry and were distributed with regard to gender and age. The objective was to have an even mix of gender, age and positive/negative attitude to the ISA-system (all drivers in the large-scale trial answered a short questionnaire with questions about, among other things, their attitude towards different ISA-systems). Forty of these were selected for this study. However, since the test-drivers had to agree to the installation of the system voluntarily there is a slight bias towards positively inclined middle-aged men in the ISA-driver population, which is reflected in the test-driver sample in this study as well.

2.2. Test route

A specific route of 33 km, which took approximately 45 min to drive, was designated for the drivers. The test route consisted of varying driving conditions including all the legal speed-limits in Sweden, both inside and outside Lund. It was divided into smaller sections with the same characteristics, e.g. one with a 30 kph speed limit with a high number of pedestrians, one with 50 kph speed limit with few pedestrians, etc. In total there were 23 different observed sections, categorized into five different street types, see Table 1.

Initially, there was a section of motorway included in the study as well, but for practical reasons the section could only stretch from one on-ramp to the nearest off-ramp, which meant that the drivers seldom reached normal motorway driving speeds. On several occasions they chose to slow down behind a truck or another slower vehicle to make sure they did not miss the off ramp. Hence, the motorway section was excluded from the study.

2.3. Apparatus

The instrumented vehicle, used for the in-car observations as well as for the observer training, was equipped with video cameras and data-loggers that registered, among other things, speed and distance to vehicle in front.

The instrumented vehicle, a Toyota Corolla model 1999 with instrumentation designed and installed by VTT, Technical Research Centre of Finland, contained three cameras, one facing forward, one facing backward and one facing the driver, see Fig. 1. The camera facing forward was used to monitor traffic in front of the vehicle. The camera facing backward showed the traffic
behind the vehicle, and the back corners of the vehicle, making it possible to study its lane positioning. In the front of the vehicle there was also a laser radar measuring the distance to the vehicle in front. Data from the in-car observations was also used in another study in the Large scale ISA-trial (Hjälmdahl, 2002).

The test drivers’ own vehicles and the instrumented vehicle were equipped with a digital map over the city of Lund in combination with a GPS positioning system and a data logger registering, among other things, time, position, speed and speed limit. The loggers in the vehicles stored data with a 5 Hz frequency. During the first month of the large-scale ISA trial the active accelerator pedal was not activated to generate data for normal driving, and it is the data from this period that has been used for speed comparisons in this study. The logged speed was calibrated by GPS to actual speed.

3. Observed variables and their validity

The variables studied according to a standardised scheme (quantitatively) during the in-car observations are: speed, speed adaptation, yielding behaviour, time gap, behaviour towards

Table 1
The different street types included in the study

<table>
<thead>
<tr>
<th>Street type</th>
<th>Speed limit (kph)</th>
<th>Driving speed (kph)</th>
<th>Number of observed sections</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arterial road</td>
<td>50</td>
<td>50</td>
<td>6</td>
<td>Mainly car traffic, signal-controlled junctions and crossings</td>
</tr>
<tr>
<td>Main street</td>
<td>50</td>
<td>50</td>
<td>5</td>
<td>Cyclists and road-side parking to some extent</td>
</tr>
<tr>
<td>Central street with mixed traffic</td>
<td>30–50</td>
<td>30</td>
<td>2</td>
<td>A large proportion of pedestrians and cyclists, interactions mid section, cars and vulnerable road users on equal terms</td>
</tr>
<tr>
<td>Rural through road</td>
<td>30–50</td>
<td>30–60</td>
<td>5</td>
<td>Rural road through small towns, few interactions</td>
</tr>
<tr>
<td>Rural road</td>
<td>70–90</td>
<td>60–100</td>
<td>5</td>
<td>Rural road</td>
</tr>
</tbody>
</table>

Fig. 1. The shooting directions of the three cameras in the instrumented vehicle.
vulnerable road users at crossings, lane usage/lane change, use of indicator and overtaking. In addition to the standardised observations, a number of variables are studied in a descriptive way (qualitatively): conflicts, interaction, communication and special events.

One of the strongest safety indicators studied in the in-car observations is conflicts, which is defined as a near-miss accident. Serious conflicts are, like traffic accidents, a result of a breakdown in the interaction of the road user, the environment and the vehicle. A serious conflict has the same development of events as an accident, with the exception that collisions rarely happen in conflicts and few if any are injured. There are between 3000 and 40,000 conflicts for each police-reported injury accident, depending on the severity and the type of the conflict. The relation between accidents and conflicts was shown by Hydén (1987). Conflict technique was taken further by Svensson (1998) who validated the relationship between traffic events according to a severity hierarchy.

Speed, naturally one of the most important safety indicators in the in-car observation method, is studied here with regard to the speed limit, the surrounding situation and to speed adaptation at junctions/obstacles. By obstacles is meant road works, parked cars, etc., that is, obstacles that require some kind of evasive action from the driver, like braking or yielding. Speed adaptation with regard to the surrounding situation means that there is no physical hindrance to be passed, but that the speed should be adapted anyway, for instance if there are children close to the vehicle or there is a bus letting off passengers. The relationship between speed and accidents is well-known and has been the object of more than 100 scientific studies throughout the world. Most of these are summarised in the Norwegian Road safety handbook (Elvik et al., 1997). The Swedish Road and Transport Research Institute has developed a model for estimating the effects of changes in mean speed on road safety (Nilsson, 2000; Nilson and Anderson, 1997). The model is:

\[
\text{Injury accident rate}_{\text{after}} = \text{Injury accident rate}_{\text{before}} \times \left(\frac{V_{\text{after}}}{V_{\text{before}}}\right)^2
\]

\[
\text{Fatal accident rate}_{\text{after}} = \text{Fatal accident rate}_{\text{before}} \times \left(\frac{V_{\text{after}}}{V_{\text{before}}}\right)^4
\]

This model has since been validated by other studies in connection with changes in speed limits on Nordic roads as well as in some other countries, for instance the National Highway Traffic Safety Association’s report from the increase of speeds on American motorways from 55 to 65 mph (1998). A more recent study on the effects of the increase of speed limits in Washington State shows that the incidence of fatal crashes more than doubled compared with what would have been expected without the increase of speed limits (Ossiander and Cummings, 2002). Salusjärvi (1981) came to a similar conclusion when he showed that the relationship between change in average speed and the number of accidents was linear. Finch et al. (1994) looked at the effects of increases or decreases in average speed had on accidents on different roads and in different countries. They found that for every 1 mph change in average speed there was an associated change of 5–9% in accidents.

Late/hard braking or bad speed adaptation at junctions/obstacles, which is another variable studied in the in-car observations, has a logical connection to safety. Nygård (1999) studied the possibilities of analysing traffic safety with the help of braking patterns from speed profiles. He found that there was a difference between conflict braking and hard braking without a conflict. Conflict braking started and ended more abruptly than normal braking. He further found that the
derivative of the acceleration (m/s\(^3\)) for normal braking had a maximum value of \(-8.0\) m/s\(^3\) whereas traffic conflict braking had a minimum value of \(-9.9\) m/s\(^3\). Hydén (1987) determined that braking is the evasive action in 93% of all conflicts and in 88% of all accidents in urban areas. A Finnish study of accident data on urban and rural roads showed that braking was the evasive action in 72.8% of the cases where there was an evasive action (Hantula, 1994). Noda et al. (1996) came to similar results when looking at acceleration noise, that is acceleration with regard to mean acceleration. They found that a high amount of acceleration noise was a good indicator of accidents, both for certain stretches of road and for individual drivers.

In an in-depth study of accidents over four years in the city of Växjö, Carlqvist and Persson (1988) identified the critical behaviour that led to these accidents. They found that too high speed in relation to the environmental and interactional demands accounted for 21% of the accidents. They further found that too small headway accounted for 5% of the accidents and erroneous yielding behaviour for 26% of the accidents. Risser (1997) found a relationship between police-reported accidents and observed behavioural variables such as exceeding the speed limit, too short headway to vehicle in front and late lane changes.

Apart from the findings of Risser and Carlquist et al. there has not been much work on the connection between headway and traffic safety; many studies, e.g. Winsum and Heino (1996), merely state that there is a natural and logical connection between the two and that the number of rear-end collisions account for a great deal of traffic accidents involving two or more vehicles. Nilson (1993) reviewed the literature regarding this issue and concluded that a direct connection between traffic safety and short headways could not be established. Mäkinen and Kulmala (1987) studied accidents and headway on three urban roads where one had a mean headway of four seconds and the other two had a mean headway of 16 s. They found that rear-end collisions were the second most frequent accident type on the road with the four second headway, whereas there were no rear-end collisions reported on the other two. In the Norwegian traffic safety handbook, Elvik et al. (1997) stated, after reviewing studies on systems which automatically controlled the distance to the vehicle in front, that the safety effect of these systems was around a 50% reduction in rear-end collisions and a overall accident reduction of 2–12%, depending on the frequency of rear-end collisions in the individual countries where the studies were carried out.

One of the advantages with the in-car observation method is that an assessment of interactions in real traffic can be made. In this study interactions are regarded as yielding behaviour at junctions, behaviour in interactions with vulnerable road users and the free observations of interactions. Carsten et al. (1989) looked at contributory factors in urban road traffic accidents and found that failure to yield was one of the main driver failures leading to accidents. Behaviour at interactions with vulnerable road users is categorized in this study as the driver yields early, the driver yields late, forces pedestrian to stop, puts pedestrian in danger and pedestrian waits at roadside, and similarly for interactions with cyclists. Drivers’ behaviour towards vulnerable road users is naturally an important safety indicator, but it is also an indicator of the drivers’ situation awareness, communication skills and, in Sweden and many other countries, his law abidance. Várhelyi (1997) found that in critical encounters, i.e. when the approaching car and pedestrian are on a collision course, 3 out of 4 drivers maintain the same speed or even accelerate and only 1 out of 4 slows down (but not necessarily stops). This clearly shows that different give-way behaviour is used as a means of communication from the driver to the pedestrian. In a study of the large-scale use of roundabouts in an urban area Hydén and Várhelyi (2002) concluded that car drivers
showed more consideration for vulnerable road users and pedestrians got priority twice as often after the introduction of roundabouts. They also showed that the number of expected injury accidents, estimated from the observed serious conflicts, decreased by 60% for bicyclists and 80% for pedestrians. Similarly, Towhiat (2001) showed that the introduction of speed humps improved drivers' give-way behaviour towards vulnerable road users, and at the same time decreased the number of serious conflicts between these groups.

Most of the variables in the in-car observations have a well documented relationship with traffic safety. There are, however, a few variables in the list at the top of this chapter that are more difficult to validate through observations and literature research. This is especially evident for communication and interaction for which it is hard to find a direct and measurable connection to traffic safety. However, their importance for traffic safety is intuitively recognized. The drivers' use of indicators is another variable that has a natural connection to safety but it is hard to find solid evidence in support of the relationship.

4. Observer training and reliability

4.1. Training

One of the aims of this study was to establish a procedure for training and evaluation of the observers. In this case three students from the civil engineering programme were selected as observers. They had attended the traffic engineering courses so they had general traffic engineering knowledge, but they were not experts.

The training consisted of three parts, each of which dealt with the free observation (qualitative) and with the coding observation (standardised).

4.1.1. Part 1: theory

After a short introduction into the principles of use of the in-car observation method—assessment of drivers, assessment of trips, assessment of infrastructure and identification of problems—information is given on what the observation method aims at studying and why. This includes both heuristic aspects—plausible reasons according to common sense and to experience—and empirical data: The latter aspect refers both to earlier research that has shown relationships between certain types of behaviour and accident risk, and to validation studies carried out in connection with the development and use of the in-car observation method within the framework of different projects.

4.1.2. Part 2: elaboration on observation variables

After the theory part the trainees themselves elaborate on possible observation variables, according to their own experience as road users. The scheme of variables that they develop in this way is then, in an iterative, process modified, corrected and completed with the help of feedback provided by the trainer(s) and general discussion. This part of the training should end with observation forms, the contents of which sufficiently resemble the original in-car observation forms.
4.1.3. Part 3: practical training

On a test route that has already been selected, or that is selected ad-hoc for the training, up to four trainees together observe the test drivers. All observers carry out the same part of the observation, either coding observation or free observation. It is very important to discuss observer-position effects since one sees different things depending on where one sits in the car during observation. During the training, position effects were counteracted by systematically varying the position of the observer-trainees. Under real conditions, the position of the free observer is on the passenger seat beside the driver, and the coding observer takes the right backseat. All training trips are followed by a discussion of all registered events. The discussions should lead to a consensus between the observers and the teacher on how the events should be registered. The training trips are repeated until a consistency in the observers’ registrations has been reached and all of the observers feel confident about the method.

4.2. Reliability of the observations

In order to be able to follow the progress of the three coding observers, the results of their observations were compared to a key that was produced after each test run by discussing the observations and using the video-recordings and the logged variables (speed, time-gap and braking) from the instrumented vehicle as support. The three observers, one or two of the teachers and in some cases, the driver, took part in the discussions. The observation sheets from each observer were then checked and graded according to the key.

When grading the observations the following system was used: each situation according to the key was worth two points, one for registering the situation and one for categorizing it correctly. Each observer was then allocated the maximum number of points according to the key. If the observers registered the situation and categorized it correctly, no points were deducted from their scores. If they registered the situation but categorized it incorrectly, one point was deducted, and if they did not notice it at all or registered something that should not have been registered, two points were deducted. Then, each observer’s score was indicated as a percentage of the maximum available points according to the key. The observations were then repeated until the results levelled out on a satisfactory level.

The free observations were more difficult to compare quantitatively. Here the condition was that all the observers had to detect the same activities and describe them in a coherent way.

The observer’s theoretical instruction took place over two days, when theory was combined with practical training on the road. Thereafter, test-runs were carried out to give the observers experience and their progress according to the key was closely monitored. Fig. 2 shows the progress of the three observers.

As Fig. 2 shows; the scores did not improve after the eleventh run where the observers’ rating was at 91% with the scoring system described above. There are some interesting features in the figure worth mentioning. On test-run no. 4 the observers performed rather poorly and one explanation for this is that the test run was carried out late on a Friday afternoon when the traffic conditions were quite intense. During a normal test run there are approximately 50 observations but during the fourth run the number of observations was 72. This was probably too much at such an early stage of their training. It is also worth mentioning that the observers had previously carried out two coding observations and one free observation on that day, which could indicate a
certain lack of attention due to fatigue. The same phenomenon can be seen for observer 3, during test run no. 10. The observer overslept that morning and performed quite badly in comparison with the other two. However, most of the errors (relative to the other observers) were made during the first half of the test-route, which was the least intense part (see Fig. 3). The importance of the observers’ attention should therefore be stressed. No more than three test-runs per day are recommended.

In the grading system used, the observations were separated into correct registrations, errors and correct situation but wrong categorization. For test-runs 11–14 the errors accounted for 7% and the correct situation but wrong categorization for 2%. This indicates that once a situation is discovered it is quite easy to categorize it, but the problem might be to discover it in the first place, Fig. 4. An objective reflection is also that those situations that were not detected by one or more of the observers had very low severity, while dangerous yielding behaviour or late braking had a 100% detection rate.
The observers were compared to each other in order to study their inter-reliability. However, this turned out to be difficult since inter-reliability for the first few test runs was so low that no model could really describe it. For test runs 11–14, however, which could be seen as model observations, it was possible to get a measure of inter-reliability. For test run eleven 40 observations were the same for all the observers, all in accordance with the key which consisted of 49 observations. The corresponding figures for test runs twelve to fourteen were: 40 out of 49, 38 out of 48 and 39 out of 51, all the unanimous observations were in accordance with the keys since there were no observations that all of them had got wrong.

5. Test of the drivers’ consistency in speed behaviour

When it comes to observers in the car there is always some scepticism about whether the test-subjects will in fact be affected and behave in a way that does not correspond with their normal driving. In this case they even drove a car other than their own, which could also have affected their behaviour.

In order to find out whether the test-drivers’ behaviour during the test drive differed from their normal driving, the speed profiles from the test drivers’ own cars were compared with the speed profiles obtained during the in-car observation. This was made possible by the fact that all the test drivers’ private cars were equipped with a data-logger that continuously registered speed and position (among other things) for one month.

Fig. 4. The distribution of registrations for the three observers.

Fig. 5. The difference in speed of test drivers when driving their own vehicle compared to when driving the instrumented vehicle, plotted against the speed of the instrumented vehicle.
For the comparison between their own car and the instrumented vehicle, spot speeds were obtained from one mid-section spot on each of the 23 observed sections. For the drivers’ own cars the mean spot speed was used, and for the instrumented vehicle the spot speed was used, since it was based on only one passing (this could be the case for their own cars as well depending on how many times they passed the individual spots).

Since data from a data logger was used, it was impossible to control for extreme values due to congestion and weather, etc. To avoid a bias from these extreme speed differences, outlier data

<table>
<thead>
<tr>
<th>Driver</th>
<th>n</th>
<th>$v_{\text{mean}}$ PC</th>
<th>Std PC</th>
<th>$v_{\text{mean}}$ IV</th>
<th>Std IV</th>
<th>Ranks $^a$</th>
<th>Z</th>
<th>Asymp. sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>12</td>
<td>38.82</td>
<td>8.67</td>
<td>41.87</td>
<td>8.31</td>
<td>5</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>94</td>
<td>6</td>
<td>50.54</td>
<td>25.33</td>
<td>50.95</td>
<td>17.69</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>111</td>
<td>7</td>
<td>40.45</td>
<td>7.23</td>
<td>39.33</td>
<td>8.37</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>117</td>
<td>5</td>
<td>47.09</td>
<td>25.86</td>
<td>45.35</td>
<td>23.04</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>152</td>
<td>5</td>
<td>43.25</td>
<td>29.04</td>
<td>46.60</td>
<td>20.09</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>1146</td>
<td>12</td>
<td>53.54</td>
<td>14.13</td>
<td>48.44</td>
<td>14.09</td>
<td>9</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>1158</td>
<td>13</td>
<td>48.33</td>
<td>12.60</td>
<td>48.62</td>
<td>11.92</td>
<td>7</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>1293</td>
<td>6</td>
<td>52.47</td>
<td>21.05</td>
<td>44.48</td>
<td>25.39</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1375</td>
<td>3</td>
<td>51.23</td>
<td>36.51</td>
<td>46.68</td>
<td>38.56</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1514</td>
<td>8</td>
<td>45.90</td>
<td>12.38</td>
<td>49.86</td>
<td>11.99</td>
<td>3</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>1523</td>
<td>1</td>
<td>11.40</td>
<td>–</td>
<td>33.70</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1633</td>
<td>16</td>
<td>49.83</td>
<td>15.11</td>
<td>43.86</td>
<td>13.28</td>
<td>14</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>1669</td>
<td>15</td>
<td>53.98</td>
<td>15.40</td>
<td>54.56</td>
<td>16.06</td>
<td>6</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>1750</td>
<td>10</td>
<td>44.57</td>
<td>12.66</td>
<td>48.07</td>
<td>14.26</td>
<td>4</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>1825</td>
<td>8</td>
<td>38.19</td>
<td>10.58</td>
<td>41.46</td>
<td>8.53</td>
<td>3</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>1837</td>
<td>9</td>
<td>47.70</td>
<td>15.51</td>
<td>49.25</td>
<td>20.04</td>
<td>4</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>1838</td>
<td>12</td>
<td>48.72</td>
<td>14.81</td>
<td>47.71</td>
<td>11.73</td>
<td>8</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>1866</td>
<td>6</td>
<td>45.19</td>
<td>8.03</td>
<td>45.95</td>
<td>10.09</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>1895</td>
<td>7</td>
<td>43.62</td>
<td>17.23</td>
<td>46.85</td>
<td>21.27</td>
<td>1</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>1909</td>
<td>7</td>
<td>46.58</td>
<td>24.24</td>
<td>42.78</td>
<td>26.66</td>
<td>3</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>2105</td>
<td>6</td>
<td>59.76</td>
<td>25.51</td>
<td>55.53</td>
<td>23.43</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2150</td>
<td>6</td>
<td>55.95</td>
<td>21.08</td>
<td>54.67</td>
<td>19.34</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2277</td>
<td>8</td>
<td>56.82</td>
<td>11.62</td>
<td>54.29</td>
<td>13.36</td>
<td>6</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2382</td>
<td>6</td>
<td>46.16</td>
<td>24.53</td>
<td>53.41</td>
<td>23.08</td>
<td>0</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>3023</td>
<td>1</td>
<td>100.40</td>
<td>–</td>
<td>85.7</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>3289</td>
<td>2</td>
<td>53.80</td>
<td>4.38</td>
<td>51.80</td>
<td>0.56</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>3592</td>
<td>5</td>
<td>52.46</td>
<td>5.79</td>
<td>54.01</td>
<td>5.07</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3721</td>
<td>10</td>
<td>48.74</td>
<td>18.87</td>
<td>51.68</td>
<td>18.02</td>
<td>4</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>3966</td>
<td>1</td>
<td>24.10</td>
<td>–</td>
<td>27.10</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>4083</td>
<td>3</td>
<td>47.33</td>
<td>12.03</td>
<td>44.23</td>
<td>8.53</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4184</td>
<td>10</td>
<td>53.36</td>
<td>20.86</td>
<td>50.75</td>
<td>17.20</td>
<td>6</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>4583</td>
<td>3</td>
<td>46.00</td>
<td>10.77</td>
<td>43.79</td>
<td>8.14</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5094</td>
<td>3</td>
<td>47.57</td>
<td>6.63</td>
<td>46.00</td>
<td>4.44</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>5186</td>
<td>6</td>
<td>43.24</td>
<td>23.02</td>
<td>44.42</td>
<td>19.50</td>
<td>4</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^a\) Note: “−” = Negative rank, that is, the speed with the instrumented vehicle was lower than the mean speed with their own cars. “+” = Positive rank. “=” = Tie.

\(^b\) \((P < 0.05)\)
outside the 99th percentile confidence interval of the mean value for each street type was removed. No particular driver or section of road was overrepresented among the outliers. In Fig. 5 the speed difference between driving with their own car and driving the instrumented vehicle is plotted against the speed of the instrumented vehicle. A negative value in Fig. 5 means that the driver drove the instrumented vehicle slower than his own car.

Possible differences between the speed values were tested pair-wise with Wilcoxon signed ranks test ($p = 0.05$) and the analysis was carried out in three different ways. First, an analysis was carried out on all the data on an aggregated level; second, each driver was analysed individually and third, the data was analysed separately for each street type.

A majority of the test drivers (34 out of 40), in their own cars, passed one or several of the 23 measured spots included in the study. The number of passages at each spot for each driver varied from 1 to 45 and the total number of passages at one single spot varied from 3 to 136.

Of the 34 drivers there were only 2 that had a significantly different speed during the in-car observation with one slower and the other one faster, see Table 2. The number of spots each driver passed with his own car is presented in the $N$ column in Table 2 and the mean speeds are based on all of these spots. Thus, a high mean speed does not necessarily represent a fast driver, but is more likely an indication that the spot speeds are taken from roads with a higher speed-limit.

For some of the drivers in the study, the number of measured spots was not high enough to make a meaningful comparison, so an analysis was carried out on an aggregated level. The analysis of all the measured spots showed no difference in the drivers' speed when driving their own cars from the speed during the in-car observation. The overall mean speed is 48.4 kph for their own cars and 48.1 kph for the instrumented vehicle ($P = 0.273$), Table 3.

When analysing the different street types individually, no difference in speeds could be found for arterial roads, main streets and rural roads. There was a difference between the speeds for central streets and for rural through roads. For central streets the speed was 3.45 kph higher when driving the instrumented vehicle, and for rural through roads it was 3.2 kph lower. However, the number of observations is the lowest for these street types, which makes the confidence intervals of the mean values on these stretches large. For central roads the confidence intervals overlap each other

### Table 3

<table>
<thead>
<tr>
<th>Street type</th>
<th>$N$ spot speeds</th>
<th>$n$ drivers</th>
<th>$V_{\text{mean}}$ PC</th>
<th>Std. PC</th>
<th>$V_{\text{mean}}$ IV</th>
<th>Std. IV</th>
<th>Ranks$^a$</th>
<th>Z</th>
<th>Asymp. sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>238</td>
<td>34</td>
<td>48.40</td>
<td>17.21</td>
<td>48.08</td>
<td>16.08</td>
<td>133</td>
<td>103</td>
<td>2</td>
</tr>
<tr>
<td>Arterial road</td>
<td>71</td>
<td>28</td>
<td>49.80</td>
<td>4.88</td>
<td>48.98</td>
<td>4.41</td>
<td>42</td>
<td>29</td>
<td>0</td>
</tr>
<tr>
<td>Main street</td>
<td>60</td>
<td>27</td>
<td>38.70</td>
<td>6.50</td>
<td>39.00</td>
<td>6.89</td>
<td>29</td>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>Central street with mixed traffic</td>
<td>33</td>
<td>23</td>
<td>25.64</td>
<td>5.46</td>
<td>29.09</td>
<td>6.98</td>
<td>14</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>Rural through road</td>
<td>32</td>
<td>10</td>
<td>50.59</td>
<td>11.01</td>
<td>47.39</td>
<td>8.97</td>
<td>24</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Rural road</td>
<td>42</td>
<td>24</td>
<td>76.11</td>
<td>12.18</td>
<td>74.96</td>
<td>11.87</td>
<td>24</td>
<td>17</td>
<td>1</td>
</tr>
</tbody>
</table>

$^a$ Note: ‘‘−’’ = Negative rank, that is, the speed with the instrumented vehicle was lower than the mean speed with their own cars. ‘‘+’’ = Positive rank. ‘‘=’’ = Tie.

$^b$ ($P < 0.05$).
by 0.8 kph and for rural through roads by 3.7 kph. More data from these street types is needed in order to say whether there is an observer effect or the difference is merely a result of the natural variation on these streets.

6. Discussion

The aim of this study was to show that the in-car observation method is a valid and reliable tool for observing driver behaviour. The method was validated previously when Risser (1985) showed that there was a correlation between observed risky behaviour and accidents and the method has been used with good results in several observational studies since, see for instance Almqvist and Nygård (1997) and Hjäml Dahl (2002). This study has added information on whether it is possible to train observers to perform observations satisfactorily; it has also added information about whether drivers’ speed choice during observations is coherent with their normal driving. Further this study looked into the observed variables’ relevance with regard to traffic safety.

This study shows that it is possible to train observers to register behavioural data in a reliable way. The observers used in this study were three students from the civil engineering programme who was given two days of theoretical and practical instruction on the in-car observation method. This was followed by in-car training combined with a discussion after each test drive. The training was repeated and the progress of the observers was carefully monitored and compared to a key which was produced for each test drive. After repeated test drives each observer had a coherence with the key of over 90%, and the result stayed on this level for the following test drives.

One of the aims of this study was to investigate the observers’ effect on the drivers. A comparison between the speeds when driving in their own car at their own pace and when driving the instrumented vehicle with observers in the car did not show any difference in speed behaviour. In this study it was possible to compare drivers’ normal driving in their own cars with driving during the in-car observations. This was possible due to the fact that all the drivers’ vehicles were equipped with a GPS positioning system and a data-logger registering speed and position. Logged data speeds were collected from 23 spots along a test route and the drivers’ speed when driving their own vehicles was compared with the speed at which they drove during the in-car observations. The analysis showed that the drivers’ speed during the in-car observations did not differ from their normal driving and this is in line with the findings of Höfner (1967). The study by Rathmayer et al. (1999) of the observers’ effect on drivers showed that the presence of observers in an instrumented vehicle lowered the mean driving speed by 1–2 kph compared with the speed when driving alone in the instrumented vehicle. They also found that repeated driving with an observer in the car had no effect on driver behaviour, which suggests that the drivers did not get used to the observers, at least not for the first 80 km of driving. Reker et al. (1993) found that, when analysing data from test trips, one should interpret speed-behaviour from the first 35 min of the observation with caution. In the present study the test-drivers drove for 15 min before the observation started, but no signs that the drivers were driving in a way which was more coherent with their normal driving in the end than in the beginning could be found. The different results from the studies discussed above show that it is difficult to define and study the driving behaviour that may be considered normal.
There is some evidence that just having a data-logger in the car is enough to change drivers’ behaviour. This has been studied by Larsson et al. (1980), who found that the installation of data-loggers in police vehicles, together with feedback based on the logged data, virtually eliminated injury accidents for these vehicles, but without feedback there was no difference. The purpose of the data-loggers in this study was not to monitor the drivers’ behaviour in the case of an accident; the data in the loggers could not be used for that purpose since it was owned by the driver. In the large-scale trial with ISA (Várhelyi et al., 2002), where the same type of data from 284 test-vehicles was used, a comparison between the logged speeds and speeds measured in the field with radar and pneumatic tubes was carried out at 12 spots in both directions. The study showed that there was no difference in speed between the logged vehicles and the rest of the vehicles in traffic.

The variables studied in the in-car observations were originally chosen from empirical data, professional experience and common sense. A review of the literature regarding the variables has shown that there is compelling evidence of the connection to safety for speed, yielding behaviour, behaviour towards vulnerable road users at crossings and conflicts. For time-gap, lane usage/lane change and overtaking behaviour the relation to safety is less well documented, but there are studies that have showed a relationship. For the variables communication, interaction and use of indicator the evidence of their importance for traffic safety is on a more intuitive level. One can therefore conclude that any observed changes detected during the in-car observations may be interpreted from a safety perspective.

Other observational methods that are commonly used, like driving simulators or instrumented vehicles, can provide much of the information that is obtainable through in-car observations, and they have proved to be valid methods as well, see e.g. Carsten et al. (1997) for validation of a driving simulator and Rathmayer et al. (1999) for validation of an instrumented vehicle. However, if one wants to study driver behaviour in a broader sense than just static measurements like speed, acceleration and headway, for instance how drivers interact and communicate with other road users, it is essential to have human observers in the vehicle.

7. Conclusion

The in-car observation method proved to be a reliable and valid method to observe driver behaviour. The study showed that observers could be trained to observe and register the correct situations in a similar way. Further, the study showed that the studied drivers’ speed choice during the in-car observations was consistent with their normal driving. A review of the literature shows that any observed changes in the studied variables are important from a safety point of view.

Further research is needed on how extreme drivers react when being studied, provided that the sample used in this study represents the normal driver. Further research is also needed on why the results on central streets and rural through roads are contradictory. Further research on the variables time-gap, lane use/lane change and use of indicators’ connection to traffic safety is also essential.

Acknowledgements

Financial support for this work was provided by KFB, the Swedish Transportation and Communications Research Board. Thanks to Ralf Risser for his valuable comments on the work.
References


