Evaluation of the SAGAT method for highly automated driving

Benjamin Franz, Jan P. Haccius, Dennis Stelzig-Krombholz, Matthias Pfromm, Michaela Kauer, Bettina Abendroth

Institute of Ergonomics and Human Factors of Technische Universität Darmstadt, Germany;

The Situation Awareness (SA) by Endsley is an often used construct to describe human’s perception and performance while driving motor vehicles. One method to assess the SA of the driver is the Situation Awareness Global Assessment Technique (SAGAT) by Endsley which is often used in driving simulators. In the last decades, the number of advanced driver assistance systems (ADAS) increased rapidly and the trend moves towards more complex semi and highly automated systems. In those systems the driver is often used as a fallback if a failure occurs and therefore the driver has to be aware of the current traffic situation. At the present time, there is no review known to the authors, whether the SAGAT-method can be used for evaluating the SA of drivers while driving semi or highly automated vehicles. To answer this question a driving simulator study was conducted at the Institute of Ergonomics and Human Factors of the Technische Universität Darmstadt. In this study, 45 participants drove the highly automated driving concept Conduct-by-Wire and a manual car on 4 days at the interval of 7 days. It shows that there are in summary only little differences between the driving modes. It can also be shown, that the relevant cues for the SAGAT method are learned by the drivers since a significant learning effect was detected. Based on these findings the usefulness of the SAGAT method is discussed and an outlook on future work is presented.

Practitioner Summary: To evaluate the SAGAT method for highly automated driving a study with 45 participants was conducted. Each participant drove manually and highly automated on 4 days at the interval of 7 days. It shows that there are only small differences comparing manually and highly automated driving. In addition, a significant learning effect regarding the SAGAT scenarios could be shown.

Keywords: SAGAT, Situation Awareness, highly automated driving, Conduct-by-Wire, pieDrive

1. Introduction

The so called Situation Awareness (SA) by Endsley (i.a. Endsley, 1995b, 2000c) is an often used construct to describe human’s perception and performance while driving motor vehicles. It is assumed, that a driver with a high SA can perform driving situations better and with less failures than a driver with a low SA (cf. Matthews, Bryant, Webb, & Harbluk, 2001). Since in the last decades, the number of advanced driver assistance systems (ADAS) increased rapidly (Vollrath, Schleicher, & Gelau, 2011) and those systems take a growing number of tasks from the driver (cf. Jones, 2002), the assessment of the SA becomes more and more important. This is especially true for the trend in the automotive development to more complex systems which combine several ADAS in one overlying system from the driver’s point of view and thereby increase the degree of automation. In those semi and highly automated vehicles, the driver is still used as a fallback if a system border is reached or malfunction occurs. Therefore it is very important for the driver to be aware of the current traffic situation.

To assess the SA different methods have been developed (Endsley, M. & Garland, D., 2000). One of these methods is the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 2000a) which is often used to rate individual ADAS in driving simulators. However, at the present time, there is no review known to the authors, whether the SAGAT-method can be used for evaluating the SA of drivers while driving highly automated vehicles.

Therefore, this article focuses on the evaluation of the SAGAT method for highly automated driving. First, the fundamentals of this article are described in detail (chapter 2). Afterwards the driving simulator study is presented, which was used to evaluate the SAGAT method (chapter 3). The article closes with the presentation and discussion of the results and an outlook on future work (chapter 4 and 5).
2. Fundamentals

In this chapter, the fundamentals of this article are described in detail. First, the construct SA and the general SAGAT method are described in detail (chapter 2.1). Afterwards, the highly automated driving concept Conduct-by-Wire is presented (chapter 2.2).

2.1 Situation Awareness and the Situation Awareness Global Assessment Technique (SAGAT)

Driving is a complex process, which is based on many individual decisions. According to Endsley (2000b), the basis for these decisions is the knowledge about the current traffic situation. These knowledge can be summarized by the term Situation Awareness (SA) which is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1988a). Accordingly, her model of SA was segmented into three levels (see Endsley, 1988a): The first level deals about the perception of relevant elements, in context of driving (e.g. the street course, road user, signs or speedometer). The second level deals with the comprehension of these elements (e.g. meaning of the sign). On the third level, the traffic situation in its entirety and its future development is described (e.g. predicted trajectory of all other relevant road users).

To assess the SA of the driver a large number of methods have been developed. One method which was originally developed for aviation (Endsley, 1987, 1988b, 1990) is the Situation Awareness Global Assessment Technique (SAGAT) (see also Endsley, 2000b). However, in the last years the SAGAT method was also used to assess the SA of drivers in automotive context (Rauch, Gradenegger, & Krüger, 2008).

The idea of the SAGAT method is to suddenly freeze and fade out the current traffic situation (Endsley, 1995a). Therefore, the method is mainly used in combination with a (driving) simulator. After the sudden fade out, the driver is asked about the traffic situation. In general, the asked questions should be unpredictable and target directly on the three levels of SA. For example, the driver could be asked to reconstruct the position of current road users (level 1), describe their predicted trajectory (level 3) and specify who has the right of way (level 2). Afterwards, the simulation is continued with the same parameters. The main advantage of the SAGAT method is described by Endsley (1995a) as its objectivity, since all answers can be matched with the actual situation.

2.2 Highly Automated Driving Concept Conduct-by-Wire

The idea of the project “Conduct-by-Wire”, which is founded by the German Research Foundation (DFG), is to transfer the driver-vehicle interaction from providing continuous control inputs with the steering wheel and pedals towards a control were the driver controls the vehicle with the help of maneuver commands (e.g. "turn left", "lane change right") and parameters ("set speed) (Schreiber, Kauer, & Bruder, 2009; Winner & Hakuli, 2006; Winner & Heuss, 2005). Those maneuvers are passed to the vehicle with the help of a new interactions concept called “pieDrive” (Franz, Kauer, Bruder, & Geyer, 2012). pieDrive consists of a contact analogue head-up-display and a touchpad input device, which is integrated in the right armrest of the driver’s seat. In the contact analogue head-up-display the current available maneuver are presented to the driver with the help of a semicircle-shaped menu (figure 1 top). Additionally, the current trajectory is presented to the driver over a contact analogue arrow lying on the road (figure 1 top). After the driver puts one finger on the surface of the touchpad the semicircle menu is automatically created around the first touched point (figure 1 bottom left). To select a maneuver the driver moves the finger towards the desired menu segment. In the head-up-display a selected maneuver is highlighted and an additional arrow previews the trajectory before executing (figure 1 middle). To execute the maneuver the finger is lifted of the surface (figure 1 right).

A more detailed description of the pieDrive concept can be found in (Franz, 2014; Franz et al., 2012). After the maneuver is passed to the vehicle, it is automatically executed by the vehicle. Therefore, the vehicle monitors surrounding traffic, recognizes traffic rules and reacts accordingly (see i.a. Franz, 2014; Franz et al., 2012; Franz et al., 2012).
3. Method

In this chapter the method of the driving simulator study is described. Therefore, the study itself (chapter 3.1), the used SAGAT scenarios (chapter 3.2) and the implementation of the SAGAT method (chapter 3.3) are presented. Afterwards, the participants of the study are shown (chapter 3.4).

3.1 Driving Simulator Study

The driving simulator study was conducted at the Institute of Ergonomics and Human Factors at Technische Universität Darmstadt between August 2013 and February 2014. The participants drove in total 8 different tracks on 4 days with the temporal distance of one week between the days (figure 2) (see also Franz, 2014). The study included 6 short distance drives, which took about 30 minutes and two longer drives with a length of app. 1.5 hours. Four short (a, c, e, g; figure 2) and one of the long drives (d; figure 2) had to be performed with Conduct-by-Wire and pieDrive. The other drives (b, f, h; figure 2) were driven manually (ordinary steering wheel and pedals with automatic transmission). To minimize impacts of learning effects, the short drives on the first (a, b), the short drives on the last day (g, h) and the two long drives (d, f) were permuted. The short drives on day 2 and day 3 were not permuted to track and describe a suspected learning effect regarding the interactions concept pieDrive.

On each short drive one SAGAT scenario was placed, whereas on each longer drive a total number of three SAGAT scenarios were prepared (figure 2). In the following chapter, the SAGAT scenarios and the implementation of the method are described in detail.
3.2 SAGAT Scenarios in the Driving Simulator Study

Within the short drives on the first and the last day (a, b, g, h) the SAGAT scenario was located in urban terrain on a 4-way intersection (figure 2 “city”). A similar scenario was also integrated in the two long drives on day 2 and 3 (figure 3 “city”). All 6 SAGAT city scenarios are based on exactly the same road layout. The scenarios differ, however, in the presented traffic signs, the available traffic and the environment to avoid the recognition of the scenario (figure 3). Within all SAGAT city scenarios the drivers drove on the priority road with a speed limit of 50km/h. In addition, there was one oncoming vehicle which changed in colour and type in the different scenarios. Furthermore, on each scenario there was one vehicle approaching the intersection from the right which stops at the stop line.

3.3 Implementation of the SAGAT method

After the simulation was stopped and the screen was switched black for the participants, the drivers had to fulfil a reconstruction task. Therefore, 3 different intersection layouts were presented to the participants from which they had to choose the correct one (figure 4). The chosen layout was then fixed on a small whiteboard. In the next task the participants had to place the traffic signs and other road users with the help of prepared magnets (figure 4 right).

After the reconstruction task the driver was asked questions about the traffic situation verbally. During this interview the reconstructed traffic situation on the whiteboard could be used to explain the situation further. Additionally, the participants had to draw the predicted trajectory of all road users and of their own vehicle (see arrow drawings on figure 4).
3.4 Participants of the Driving Simulator Study

The study involved a total number of 45 participants (22 female) with an average age of 22.3 years (SD: 1.96). At the time of the study, all participants had a valid driver license of the Federal Republic of Germany for passenger cars (license held in years: \( x = 4.4 \); SD = 2.1).

4. Results

In this chapter, the results of the driving simulator study are presented. As previously described, the results should determine whether the SAGAT method can be used to assess the SA while driving highly automated. Therefore, the differences between driving manually and with Conduct-by-Wire will be shown only if findings on the SAGAT method can be made. In addition, no global SAGAT score was built. For the analysis of the results, the answers of the participants during the reconstruction task and the interview were compared to the data recorded by the simulator. First, the answers were separated in different categories and afterwards the categories were used to compare the result with the help of contingency tables (two-way tables) and the Chi-squared test (Yates, 1934). Below, the results are presented in the chapter "layout of the intersection" (chapter 4.1), "perception of other cars" (chapter 4.2) and "right of way" (chapter 4.3). Additionally, learning effects are described in chapter 4.4.

4.1 Layout of the Intersection

There were three different possibilities for the participant while choosing the intersection layout (figure 4): t-crossing, 4-way-intersection without stop lines on two roads (correct layout); 3: 4-way-intersection with stop lines on all roads (see figure 5). While driving highly automated, 76.5% of the participants chose the correct layout (2). Similarly, 76.4% participants chose the correct while driving manually. The 4-way-intersection with stop lines on all roads (3) was chosen by 11.4% (highly automated) respectively 9.8% (manual driving). The distribution of the answers can be found in table 1. The chi-square test of the corresponding contingency table shows no significant difference between the two driving concepts (\( p = 0.862 \)).
Table 1. Distribution of the answers regarding the layout of the intersection.

<table>
<thead>
<tr>
<th>layout of the intersection</th>
<th>4-way-intersection without stop lines on two roads</th>
<th>4-way-intersection with stop lines on all roads</th>
<th>t-crossing</th>
</tr>
</thead>
<tbody>
<tr>
<td>highly automated (Conduct-by-Wire and pieDrive)</td>
<td>76.5% (101)</td>
<td>11.4% (15)</td>
<td>12.1% (16)</td>
</tr>
<tr>
<td>manual driving</td>
<td>76.4% (94)</td>
<td>9.8% (12)</td>
<td>13.8% (17)</td>
</tr>
</tbody>
</table>

4.2 Perception of other Cars

To rate the perception of other cars, a score was developed, which describes the failures while placing the vehicles in the reconstruction task. Therefore, missing, misplaced and placed fictive vehicles were counted and summarized. Accordingly, a score of 0 represents that no mistake has happened (both vehicles were placed at the right position and no other vehicle was positioned). The results show that especially 2 and 3 mistakes occurred more often while driving highly automated (see table 2). Additionally, the average score while driving manually is significantly lower compared to driving highly automated (0.27 to 0.49; p = 0.034; Wilcoxon signed rank test). However, in total both modes showed over 70% completely correct reconstruction of the other road users.

Table 2. Number of mistakes in perception of other cars.

<table>
<thead>
<tr>
<th>number of mistakes</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>highly automated (Conduct-by-Wire and pieDrive)</td>
<td>71.2% (94)</td>
<td>12.1% (16)</td>
<td>12.9% (17)</td>
<td>3.8% (5)</td>
</tr>
<tr>
<td>manual driving</td>
<td>77.2% (95)</td>
<td>18.7% (23)</td>
<td>4.1% (5)</td>
<td>0% (0)</td>
</tr>
</tbody>
</table>

4.3 Right of Way

The results of the reconstruction task and the interview regarding the understanding of the driver for the right of way can be found in table 3. It shows that in less than 5% of the drives all signs were placed completely correct by the participants (priority road and stop signs were placed). However, the right of way was correctly recognized by more than 80% of the participants (priority road, stop or give way signs were placed in different combinations). The chi-square test of the corresponding contingency table shows no significant difference between the two driving modes (p = 0.922). By comparing the results of the reconstruction task with the interview it can be noticed that more correct answers regarding the right of way were given by the interview (see table 3).

Table 3. Distribution of the answers regarding the right of way.

<table>
<thead>
<tr>
<th>driving mode</th>
<th>correct answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>interview</td>
<td>manual driving</td>
</tr>
<tr>
<td></td>
<td>highly automated (Conduct-by-Wire and pieDrive)</td>
</tr>
<tr>
<td>reconstruction task</td>
<td>manual driving</td>
</tr>
<tr>
<td></td>
<td>highly automated (Conduct-by-Wire and pieDrive)</td>
</tr>
</tbody>
</table>
4.4 Learning Effects

Looking at the results over the test duration shows time based differences between the drives. In this section this effect is presented on the example of the failure score to assess the perception of other cars described in section 4.2 (see table 4). It can be seen that significantly ($p = 0.00$) more failures were made on the first day of the study compared to the last day. Additionally, it shows that during the permuted drives on the first day (a, b) significantly more failures occurred in the reconstruction task while driving highly automated ($p = 0.007$). This difference cannot be shown ($p = 0.083$) on the last day of the study (g, h).

Table 4. Learning effects on the failure score.

<table>
<thead>
<tr>
<th></th>
<th>Highly automated</th>
<th></th>
<th>Manual Driving</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>average failure score</td>
<td></td>
<td>average failure score</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>1.09</td>
<td>b</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>0.30</td>
<td>f</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>0.07</td>
<td>h</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

5. Discussion and Outlook on Future Work

Neglecting the learning effect, the evaluation of the driving simulator study showed only minor differences between the results of the two tested driving modes. Therefore the SA of the driver would be rated similar while driving manually compared to driving highly automated. Only small differences can be seen e.g. when examining the number of mistakes in car perception. Here, the number of errors was slightly higher for highly automated driving. Therefore, it would be reasonable to assume the same or almost equal SA for both modes. But as can be seen in table 3 SAGAT reveals less information than an interview does when concentrating on the right of way. This leads to the assumption that not all of the knowledge a driver has, can be examined with the help of SAGAT. Knowing this, the meaningfulness of the SAGAT results has to be questioned.

Considering the learning effects, it turned out that the average number of errors decreases with increased familiarity with the method. This leads to two interesting conclusions: Firstly, the SA in the highly automated driving mode increased over time. Secondly, the usage of SAGAT has to be trained (as recommended by Endsley). To examine whether the SA really increased over time an eye-tracking analysis was conducted that compared eye movements and gaze durations between manual driving and highly automated driving. A first investigation with data of 15 of the 45 participants showed that the percentage of gazes in the considered areas of interest (road, environment, car interior) was equal for both modes in the beginning and then differed significantly for the last measurement point (Franz, 2014). It turned out that the number of gazes on the road decreased over time for highly automated driving (Franz, 2014). Therefore, eye-tracking analyses and SAGAT show different results concerning the amount and kind of information that is processed by the driver. Combined with the decreased number of errors over time this leads to the assumption that drivers learned which cues they have to track to be successful in the SAGAT task, but did not maintain the driving behavior they showed during manual driving.

Overall, it is questionable how meaningful the results of SAGAT are, especially for – but not restricted to – highly automated driving. The fact that drivers have to train the method even if they are successful drivers in the real world (and therefore detect and process the necessary information) leads to the assumption that normally the information is not stored explicitly. This is confirmed by the fact that the interview revealed more information than SAGAT did. Combined with the fact that eye-tracking and SAGAT weren’t compliant, further investigations are needed to assess the quality of SAGAT for highly automated driving.

In a first step the remaining eye-tracking data will be analyzed and compared with the SAGAT results. Following, further studies that use additional methods to assess the Situation Awareness will be conducted. But above all that stands the still open question: How much SA do you need to drive successfully? And is the amount of SA and the kind of information you need the same for manual driving and highly automated driving? Before having answered these questions – perhaps on a theoretical basis – a statement about the suitability of any method is hardly possible.
References


