Retrieving vehicle control after automation:

A comparison between two methods to measure situation awareness



Bachelor thesis 2013

Author:	Tom Hueting	
1 st Supervisor:	prof. dr. ing. Willem Verwey	
2 nd Supervisor:	ir. Arie Paul van den Beukel	

University of Twente

Faculty of Behavioural sciences

Study of Psychology

Retrieving vehicle control after automation:

A comparison between two methods to measure situation awareness

Bachelor thesis

Enschede, February 25th, 2013

Author:	Tom Hueting	
1 st Supervisor:	prof. dr. ing. Willem Verwey	
2 nd Supervisor:	ir. Arie Paul van den Beukel	

University of Twente, Enschede Faculty of behavioural sciences Study of Psychology Theme: Human Factors and Media Psychology

Abstract English

Automation in vehicles is becoming more advanced and more commonly available in consumer cars. One recent development is the ability of cars to drive themselves in traffic jams. This has consequences for the way people deal with, and operate their cars. Higher levels of automation influence the level of involvement in the driving task. This can lead to a lower level of situation awareness (SA), which in turn leads to poorer driving performance.

Within this research several methods for measuring SA were investigated in order to determine which measure is best suitable for evaluating driver-vehicle interfaces in automatically operated vehicles. Therefore a simulator experiment was conducted in which participants had to respond to a critical situation after they were brought out of the loop, i.e. without any situation awareness. It was then analyzed whether the criticality of the situation had any influence on driving performance and level of SA. Furthermore, the SA-scores provided by the Situation Awareness Global Assessment Technique (SAGAT), and the Situation Awareness Rating Technique (SART) were evaluated.

It was found that criticality did have an effect on driving performance. People had significantly less accidents as they had more time to respond to the situation. Furthermore it was found that only the SART questionnaire correlated with criticality of the situation and number of successful take-overs.

Samenvatting Nederlands

Automatiseringen in voertuigen worden steeds geavanceerder en worden steeds algemener in personenvoertuigen. Een recente ontwikkeling binnen de auto-industrie is de mogelijkheid van auto's om zichzelf te besturen in files. Dit heeft consequenties voor de manier waarop mensen hun auto's besturen en hoe zij met de systemen omgaan. Hogere niveaus van automatisering beïnvloeden de mate van betrokkenheid bij de rijtaak. Dit kan leiden tot lagere niveaus van 'situation awareness' (SA), wat weer leidt tot vermindering van de rijprestatie.

Binnen dit onderzoek zijn verschillende methoden voor het meten van SA belicht met als doel te bepalen welke methode het meest geschikt is voor het evalueren van verschillende bestuurder-voertuig interfaces in automatisch bestuurde auto's. Daartoe is een simulatorexperiment opgezet waarbij proefpersonen moesten reageren op een kritische situatie nadat zij 'out of the loop' waren gebracht. 'Out of the loop' wil zeggen dat zij zich niet bewust waren van de omgeving van de auto (geen SA hadden). Daarna werd geanalyseerd of the kritikaliteit van de situatie van invloed was op de rijprestatie en de mate van SA. Verder werden de SA-scores resulterend uit de 'Situation Awareness Global Assessment Technique' (SAGAT) en de 'Situation Awareness Rating Technique' (SART) geëvalueerd.

Er werd gevonden dat de kritikaliteit van de situatie een effect had op de rijprestatie. Mensen hadden significant minder ongelukken naarmate zij meer tijd hadden om te reageren op de situatie. Verder werd gevonden dat alleen de SART vragenlijst correleerde met de kritikaliteit van de situatie en het aantal succesvolle overnames.

Contents

1. In	ntroduction	11
1.1	Interaction with automated vehicles	12
1.2	Situation awareness	13
1.3	Methods for measuring situation awareness	14
1.5	Evaluation why SAGAT and SART were chosen	16
The pre	esent experiment	17
2. In	ntroduction	17
3. N	1ethod	19
3.1	Participants	19
3.2	Procedure	19
3.3	Task	20
3.4	Apparatus	21
4. R	esults	22
4.1	Effect of criticality on number of successful take-overs	
4.2	Relation between situation awareness and criticality of take-over	
4.3	Relationship between SA-scores and percentage of successful take-overs	25
5. D	Discussion	25
5.1	The relation between successfulness and criticality of take-over	25
5.2	Relation between situation awareness and criticality of take-over	
5.3	Relation between situation awareness and successfulness of take-over	
5.4	Suitability of methods for measuring situation awareness	
6. R	ecommendations	
7. R	eferences	
8. A	ppendixes	30
Inde	x of Appendixes:	30
Appe	endix 1	
Appe	endix 2	

1. Introduction

In the past years the automotive industry has been undergoing changes towards more automated forms of driving. Several technological advances have been made, one of these advances is traffic assistance, the ability of a car to drive itself during traffic jams on relatively straight roads and below a specified speed. This could offer more comfort for the driver as well as improve traffic efficiency. With this technology becoming more popular, questions arise about the ability of the human operator to cope with this different way of controlling a vehicle. With increasing automation the role of the driver changes from actively operating the vehicle to passively monitoring the system. This change causes out-of-the-loop performance problems, meaning that the driver is no longer involved with the driving task at hand, causing him to lose overview of the situation, that is, a loss of situation awareness (Endsley, 1996). This is especially critical when the driver needs to take over control of the vehicle from the automated system, e.g. when the system fails or, in case of traffic assistance, when the speed exceeds the limit of operation. System failures can occur during several common situations on the road, for example as a result of absent road markings, sudden bad weather conditions, accidents, road works or merging lanes. It has been found that a driver who is out of the loop is slower to recognise an event in his surroundings (Merat, 2009), therefore a warning message given by the system will take longer to be processed. This is especially problematic in a situation where time is critical, e.g. when an emergency take-over is requested by the system.

Due to the slower recognition time it is important that the driver-vehicle interface is designed in such a way that it can give a warning message sufficiently early so that the driver is still able to bring himself back into the loop to handle the situation manually. In order to be able to do this the driver must be aware of the situation. Situation awareness is a frequently used term in automation studies and has been frequently researched. However in combination with operating a vehicle there are still questions remaining. In a recent presentation regarding the challenges in automated driving, summarising the current efforts of many institutions, it came forward that investigation is needed on the amount of time needed for re-involvement of a driver who is out of the loop (Hoeger, 2012). This issue will be addressed in the experiment.

A second issue that will be addressed is how to assess if situation awareness has been regained. This can be determined through the use of several measures which will be discussed in the paragraph 'Methods for measuring situation awareness'.

The answers to these issues can be used to choose a method for measuring situation awareness, and to develop an assessment framework which can be used to assess different varieties of a vehicle-driver interface with respect to their offered support in regaining situation awareness.

1.1 Interaction with automated vehicles

Automation in vehicles has seen several great developments throughout the last several decades. Examples are adaptive cruise control and advanced radar-based warning systems. More recently time has been devoted to the development of systems which, once activated, can control a vehicle without human interference. One of these systems is called Traffic Jam Assist (TJA). Each of the different systems for assisted driving which have been developed offer a certain level of automation. Level of automation (LOA) is a term proposed to indicate which parts of a certain task are performed by the operator, and which parts are being taken over by an computer (Endsley, 1999). TJA is a system that offers a highly automated form of driving below a speed of 60 kph. It takes over lateral and longitudinal control of the vehicle allowing it to stay in lane and follow a car in front, although for special operations like changing lanes or exiting the motorway human input is still required.

Research has shown that when the level of automation increases, the level of involvement of a person performing a certain task decreases (Endsley, 1999; Wiener, 1980). The consequence of operating a highly automated vehicle is that the driver is no longer involved in operating the vehicle. The driver's role has changed from driver of a car, to operator of a system. This causes the driver to be 'out of the loop' with regard to the driving task. Being out of the loop has several

consequences for a person's behaviour and state of mind. It may lead to overreliance on the automatic system, it can cause erratic levels of mental workload, skill degradation, and can cause a reduced level of situation awareness (Saffarian, 2012). Whereas some of these consequences are long term, reduced situation awareness is something that can already occur during first use. This is a reason why situation awareness is of great interest to the developers of automated systems.

1.2 Situation awareness

Situation awareness (SA) is a term used to describe the level of a person's awareness of his surroundings, and how his actions will impact it. Surroundings in this case can be broadly defined as the area in and around a vehicle, the state of the vehicle itself, and the way all elements of the surroundings are interacting and developing over time.

SA has been researched extensively and it has been determined that it consists of three levels (Endsley, 1996). All three levels must be achieved in order to be able to conclude that a person is aware. The first level entails perception of the different elements of the situation. The second level requires comprehension of the perceived elements. These two must then be integrated into a holistic picture of the situation so that the meaning of the individual elements can be comprehended. This level of comprehension is defined as the third level of situation awareness. Achieving the third level means that one can make predictions about future changes or behaviour of the various elements of the situation (Endsley, 1996).

Not being aware of the situation is detrimental to operating a vehicle. It should therefore be kept in mind that automating certain functions of a vehicle has a negative effect on a person's ability to take over control when the system fails. Previous research using driving simulators has shown that drivers operating a simulated automatic vehicle were significantly slower to respond to critical events which required manual intervention (Merat, 2009). This slow response is the main area of interest of the current experiment.

1.3 Methods for measuring situation awareness

Through the many years that situation awareness has been of interest, many different methods have been developed to measure it. Generally there are two types of measures; objective and subjective ones. The most commonly used objective measure is the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1988, 1989, 1995). The most popular subjective measure is the Situation Awareness Rating Technique (SART) (Taylor, 1990). Another way of determining SA in a vehicle is by observing driving behaviour. Driving behaviour entails things like steering motion and use of the pedals.

Considerable research has been done on the different methods of measuring SA. For instance during the development of the SAGAT questionnaire it was found that subjective ratings are ineffective, as the answers given by the pilot were greatly biased by the outcome of the mission. A good outcome of a mission resulted in a high rating of SA, regardless whether it was through good SA or just good luck. The answers were also affected by the amount of time that passed between the event and when the question was asked, and by other events that passed after the event of interest (Endsley, 1999). This is the reason she proposed an objective measure to avoid any false results through inadequate self-reports. In addition it was found that physiological measures such as EEG and eye-tracking provide some indication of whether information was cognitively registered. A problem however is that it is impossible to determine how much information actually remains in memory, whether the information is registered correctly, and the degree of comprehension of the registered elements of the situation.

The Situation Awareness Global Assessment Technique, developed in 1988, is one of the first validated measures for measuring situation awareness (Endsley, 1988). It has been originally developed for the aircraft industry in order to determine whether certain interface designs were successful in establishing and maintaining situation awareness. The SAGAT questionnaire proposes a solution to the poor self-reports and the dependence on good memory to answer any questions by asking questions during the execution of the task. A drawback of earlier approaches

which required the operator to answer questions during the task is that the pilot could direct attention to the requested information, falsely increasing the degree of SA reported. The SAGAT questionnaire aims to overcome this drawback. It is based on the principle of stopping simulation and blanking the screen during the execution of a particular simulated task. This allows the experimenter to ask situation-specific questions shortly after the event happened. The idea is that the amount of time between the mental event and the moment of recall is only very short, therefore reducing the influence that memory might have on the recall. The questions asked during admission of the SAGAT questionnaire are all related to key events or objects within a situation that are of immediate importance. While driving for instance, a key event could be the vehicle's speed, and a less important event is the colour of an oncoming car. The importance of events are all situation and task-specific therefore the questionnaire needs to be adjusted to meet the specific requirements for each experiment. A drawback of the SAGAT questionnaire is that it requires a simulator which can be paused. Such simulators are often of considerable cost. Furthermore a consideration with SAGAT is that it could interfere with the performance on the primary task.

The Situation Awareness Rating Technique, developed in 1990 is a subjective method for measuring situation awareness (Taylor, 1990). The participant is required to rate on a scale of 1 (low) to 7 (high) a set of 10 questions.. These questions are clustered and the answers added up into three domains; Demand, Supply, and Understanding. With a simple calculation using the total scores of each of the three domains a SA-rating is obtained. The formula is "Understanding - (Demand - Supply)". A disadvantage of using the SART is that people's answers can be influenced by their performance on the task when reporting their answers. For instance a very good take-over may lead a participant to believe his level of SA ought to have been good as well as a result. Furthermore, people are not always able to accurately assess their level of SA since they may not be aware of what they did not see. An advantage of this method though, is that it is less dependent on good memory than for instance the SAGAT. Furthermore since the test is not tailor

made for specific tasks, results can be easily compared between different tasks. This makes it simple to use, cheap, and applicable in many different situations.

There is also a shortened version of the SART questionnaire called the Three Dimensional SART (3D-SART). This is a condensed version of the normal SART which asks participants to answer only 3 questions instead of the regular 10 (see appendix 2). Each of the three questions is about one of the three dimensions the SART aims to measure (Strybel, 2007). This version is especially advantageous when it is desirable to complete the questionnaire in a short amount of time.

Lastly, there are also several objective car-specific measures related to driving performance. These measures have been found to indicate the level of situation awareness, the first measure of which is steering behaviour (Rauch, 2009; Verwey, 1999). On the basis of the frequency of steering corrections, the angle, and the speed of movement of the steering wheel, conclusions can be drawn regarding the level of SA. It has been found that steering corrections of lesser frequency, but of larger angle indicate the driver is less situation aware. Additionally the 'time to line-crossing, gives an indication of how well the driver is able to follow the curvature of the road (Rauch, 2009; Verwey, 1999). As the name suggests, TLC indicates the time remaining until the vehicle crosses one of the road markings when the current course is maintained.

1.5 Evaluation why SAGAT and SART were chosen

The experiment was conducted to evaluate the best method of measuring SA in a simulated driving environment. The choice was made that both an objective and subjective method should be evaluated. Of all objective methods the SAGAT questionnaire is used most commonly, and of the subjective methods the SART is used most often (Endsley, 1999; Salmon, 2006). Considering the methods are both found to be valid and reliable for determining the degree of SA (Endsley, 1999; Salmon, 2006) and in addition are both suitable for use in the set-up of the current experiment, it was chosen to use these two for the current experiment. Due to the fact that the level of situation

awareness was the main focus of the experiment it was chosen to not include any performance measures since it is unclear to what extend these indicate the level of SA rather than driving skills.

The present experiment

2. Introduction

An experiment was conducted to answer research questions that came forth from earlier literature, and to test the corresponding hypotheses.

Previous literature has shown that an operator of a system, who is out of the loop, is slower to respond to events in his surroundings (Merat, 2009). This raised the question of how much time would actually be needed to respond to certain events. In this case the relevant topic was how quickly a driver could be brought back into the loop to take over manual control of a vehicle. We suggested that the successfulness of take-overs would correspond with the criticality of the situation. By criticality the time is meant between simulated failure of the automated control and a collision which was varied in the experiment. We expected that as the criticality decreases, i.e. more time until a collision, the average success rate of take-over would increase.

To gain situation awareness, Endsley proposed that a person needs to pass through three stages (Endsley, 1996). This requires a certain amount of time, so the question of interest was how long it would take to regain situation awareness. In this experiment the focus was mainly on attempting to find a relation between the amount of time available to respond to a situation, and the level of situation awareness. We suggested that the level of SA would be higher in the least critical conditions. If comparing between SART and SAGAT, we expected in particular a positive correlation between SA and criticality based on the SAGAT questionnaire, due to the expected inaccuracy of self-assessment involved in SART. Besides, within various validation studies the

SAGAT technique has proven to be superior in terms of reliability, validity and sensitivity when compared to SART (Salmon, 2006).

It has also been found that situation awareness is needed to effectively and accurately respond to a certain situation. Therefore the question was whether it would be found that situation awareness in our experiment is a prerequisite of good driving performance. Based on the assumption that situation awareness is indeed a prerequisite, we expected a positive correlation between SA and successfulness of take-over.

The research questions were investigated by means of a simulator experiment. At the onset of the simulation the car was driving automatically at which point the participants performed a secondary task to ensure they were unaware of the simulated environment and the state of the vehicle (out of the loop). After a cue they had to be re-involved in the driving task to avoid an accident by taking over manual control of the vehicle. The time available to respond to the situation was varied according to three different criticalities. In the first condition, participants had 0,5 seconds after the cue to avoid an accident. The time to avoid an accident in the other conditions was respectively 1,0 and 1,5 seconds. The experiment was set up using a 2x4design with 1 repetition. A more detailed description of the task will be provided in the section 3.2.

Situation awareness was assessed through a SAGAT and SART questionnaire, as well as through the successfulness of the take-over. Based on previous literature it was assumed that the SAGAT questionnaire would give better results than the SART. This is due to the fact that it has been found that people are generally poor at reporting their own level of situation awareness

The answers to the research questions were used to evaluate the successfulness of the SAGAT and SART questionnaires in measuring SA in this setting.

3. Method

3.1 Participants

The participants consisted of 34 undergraduate students, of which 20 were female (58,8%). The mean age of the participants was 21,35 ranging from 18 till 32.

3.2 Procedure

Firstly the participants were asked to fill in a form with several demographic and other general questions (appendix 1). After completing the form they were instructed to take place in the simulator and to adjust their seat to their length. After receiving instructions about the nature and procedure of the experiment, described in the section below, the experiment started with a practise run to allow the participants to get used to the simulator.

The experiment itself consisted of 16 trials according to a 2x4x2 design, where the participants would be 'launched' onto the highway at a speed of 60kph (16,7 m/s) at which point the car would be driving automatically. After a certain period of time an alarm could sound, meaning that the automatic control had failed and a manual over-take was required to prevent an accident.

During each trial the amount and distribution of the traffic was kept the same. This is important to ensure each participant gets the same chance of resolving the situation in each trial.

In total the participants were presented with 4 different conditions which were all presented four times. The four conditions were made up of three different criticalities and one additional condition where no over-take was required. Each condition was presented twice while driving on the right lane, and twice while driving on the left lane. The conditions were presented in a different random order for each participant. This was done to ensure any possible learning effect would not occur on the same conditions for every participant.

The three criticalities describe the time participants had between the sounding of the alarm, and the moment an accident would happen if no action was taken by the driver. The amount of time varied between 0.5, 1.0, and 1.5 seconds.

19

It was chosen to alternate between starting lane to make the situation more ambiguous for the participant. It was believed that if all conditions would be presented on the right lane participants would automatically start to evade to the hard shoulder, alternating between lanes is believed to prevent this. The 'no action' condition was included in an attempt to make participants less expectant of an over-take.

At the onset of the simulated drive, when the car was driving automatically, the participants were instructed to perform the secondary task and not look at the screen in front, which presented the simulated environment. This was done to simulate a driver being out-of-the-loop. After a period of 15, 20, or 25 seconds, the automatic control failed and an alarm sounded, cueing the participants that they had to take over manual control. At this point the participants had to cease performing the secondary task and immediately take over control of the vehicle to avoid an accident. Once the situation was resolved the simulation was stopped, the screens blanked, followed firstly by the SAGAT, and then by the SART questionnaires.

During the experiment one camera faced the participants. The image was displayed on a separate screen allowing the experimenter to observe their actions. This was primarily used to determine whether the participants were looking in the mirrors during the take-over, but also to determine if they were properly concentrating on the secondary task. A second camera was facing the simulator screen, this was used to record what happened during the experiment. This could be used as reference to the experimenter about the other vehicles on the road, and to review the answers participants gave on the SAGAT questionnaire.

3.3 Task

Task instruction included mentioning the participants were going to be 'launched' onto the highway for a total of 16 times at a speed of 60kph, at which time the car would be driving automatically and the participant would be required to perform the secondary task and not look at the road. The secondary task was to play a simple game on a touch-screen. The player had to

ensure that a blue ball did not touch any of the moving red balls. This could be achieved by simply dragging the blue ball over the screen using a finger. While playing the game the participant had to bend over a little to reach the screen, ensuring the simulator screen was not visible. Figure 1 shows a screenshot and a picture of the game in use.

If was further explained that after a certain amount of time if could occur that the automatic control of the vehicle fails, at which time the alarm would sound. It was made clear that the main goal of the participant would be to avoid an accident, preferably by evading the obstacle but

any action was allowed as long as they would remain on the road. They were told that after each trial where they had to take over control of the vehicle, several questions would be asked

concerning the type and colour of the vehicles around them (SAGAT), and they were required to verbally reply. After the SAGAT they were asked to fill in the SART questionnaire provided inside the simulator.

3.4 Apparatus

The experiment was conducted using a driving simulator, a picture of which is shown in figure . The simulator is made up of three screens, positioned at a distance of approximately 200cm. This viewing distance provides a viewing angle of approximately 200°.



Figure 1. Secondary task; screenshot (above), in use (below)



Figure 2. Simulator set-up; experimenter view (above), inside view of simulator and secondary task (below).

4. Results

4.1 Effect of criticality on number of successful take-overs

To test the hypothesis that driving performance corresponds with the time available to respond to a situation, the mean success-rates for each time condition were compared. The means indicated that when the criticality decreases the success-rate increases. To test this hypothesis an analysis of variance (ANOVA) was performed on the mean success-rates. This analysis indicated that the means for each of the three conditions differ significantly (F(2,89)=20.417, p<0.001). This indicated a possible effect of criticality on success-rate.

To determine such a relation a Pearson correlation coefficient was calculated for the relation between successfulness of the take-over and the criticality of the situation. Results showed a significant correlation between the amount of successful takeovers and criticality (r = 0.541, p < 0.001).

Figure 3 illustrates the number of takeovers which were not successful for each of the three time conditions. In the most critical situation the number of unsuccessful takeovers was 47,5%. In the 1,0 second condition the percentage of accidents decreased to a little under 21%, a significant 50% decrease (F(1,59)=44.639, p<0,001). In the 1,5 second condition the number of unsuccessful take-overs of



(F(1,59)=44.639, p<0,001). In the 1,5 second *Figure 3*. Percentage of accidents avoided condition the number of unsuccessful take-overs decreased yet a little more to 12,5%, another 10% decrease compared to the second condition, but this decrease, however, was not significant (F(1,59)=2.320, p=.133).

4.2 Relation between situation awareness and criticality of take-over

Figure 4 shows a box plot displaying the SA scores provided by the SAGAT and SART questionnaires. The scores of the SART have been corrected to match the range of the scores of the SAGAT, allowing the comparison of both in one graph. The correction was done by multiplying the SART scores by 0,75 ensuring a range of 15 (from -3,5 to 11,5). The scores were then increased by adding 3,5 to make the



Figure 4. Situation Awareness by SAGAT and SART questionnaires for each condition

scores range from 0 to 15 like the SAGAT-scores. The figure shows that the overall scores resulting from the SAGAT questionnaire were higher than those that resulted from the SART. Further analyses were performed without any correction to avoid unnecessary inaccuracies.

The results from the SAGAT questionnaire were contrary to the expectation that situation awareness would increase as the criticality decreases. Instead, according to this measure, situation awareness increases slightly as the criticality decreases from 0.5 to 1.0 seconds (mean score from 8,38 tot 8,82), but then decreases slightly as the criticality decreases from 1.0 to 1.5 seconds (mean score from 8,82 to 7,41).

The SART questionnaire provided results that were in line with our expectations. Table 3 shows the mean scores for each of the three time conditions. The mean situation awareness scores increased as the criticality of the situation decreased.

An ANOVA analysis of variance shows that within both the SAGAT and SART measures several scores were significantly different for each condition, indicating some effect. Especially the SART showed a significant increase in scores as shown in table 1.

Table 1.

Condition	df	F	p
0,5 second $-1,0$ second			
SAGAT	1,67	0.669	.416
SART	1, 67	5.441	.023
1,0 second – 1,5 second			
SAGAT	1, 67	5.996	.017
SART	1, 67	0.527	.470
0,5 second – 1,5 second			
SAGAT	1, 67	3.223	.077
SART	1, 67	8.547	.005
Overall			
SAGAT	2, 101	3.391	.038
SART	2, 101	4.721	.011

Analysis of variance between mean SA scores

Note. Significant values are highlighted in bold

A paired samples t-test was performed to determine if the SAGAT and SART questionnaires provided a similar level of situation awareness. The result was that the SAGAT score for SA was significantly higher than the score provided by the SART (t(df=101)=12.295, p<0.001). However, a significant positive correlation between the two measures was found of r=.259 at a significance of p=0.008.

A Pearson correlation analysis was performed to determine if there was a relation between criticality and the mean SAGAT- and SART-scores for each of the three conditions. The result was an insignificant negative correlation between criticality and mean SAGAT scores (r=-.169, p=0.089). However, a significant correlation was found between criticality and the mean SART-scores (r=.284, p=0.004), indicating that a lower level of criticality coincides with a higher SART score.

4.3 Relationship between SA-scores and percentage of successful take-overs

We expected a positive correlation between situation awareness scores and the average successfulness of take-overs. The results showed that only the SART questionnaire was positively correlated with successfulness of take-over (r = .323 p =0.002). It was also clear that the means of the SART-scores for each of the three timeconditions increased as the criticality decreased



Figure 5. Situation Awareness by SAGAT and SART questionnaires for each success percentage

(0,5 second: 4.02, 1,0 second: 5.12, 1,5 second: 5.50) The SAGAT questionnaire showed no such correlation (r = .020, p = .852).

As mentioned earlier, both the SAGAT and SART questionnaires showed significantly different mean scores for each of the three conditions, indicating an influence of criticality on SA-scores (figure 5). It can however be seen that the SA-score provided by the SART questionnaire showed a clear positive correlation with the mean percentage of succesful take-overs for each condition.

5. Discussion

5.1 The relation between successfulness and criticality of take-over

One of the research questions was to examine the relationship between how much time was available to respond to a situation (criticality), and the degree in which participants were able to successfully deal with the situation, i.e. avoid an accident.

From the results we can conclude that different levels of criticality were of influence on the driving performance of the participants. As the criticality decreased, the amount of successful take-

overs increased. Meaning that although criticality was high in all conditions, the less critical conditions were more manageable to the participants than the more critical ones.

5.2 Relation between situation awareness and criticality of take-over

We wanted to determine if there is a relation between the amount of time that was available to respond to a situation, and the scores provided by the different measures for measuring situation awareness.

The data indicates that the SA-score provided by the SART questionnaire is correlating with the criticality of the conditions, as was expected. In addition it can be seen that the level of SA provided by the SART increased as the criticality decreased. This indicates that participants in the less critical conditions were better able to divide their attention between observing the traffic and controlling their own vehicle. The SAGAT questionnaire did not provide the same results, there was no correlation with the number of successful take-overs

It was shown that the level of situation awareness resulting from the SAGAT and SART questionnaires were significantly different from each other. This means that at least one of the measures is providing a false level of SA, although is it also possible that both are inaccurate to some degree.

5.3 Relation between situation awareness and successfulness of take-over

After examining the relation between SA and criticality, we also wanted to find out if there was a relation between SA scores and the actual number of successful take-overs.

The SART questionnaire has indicated such a relation, by showing some positive correlation with the successfulness of take-over. This means it is likely that we can assume that when SA increases, so does the chance for a successful take-over.

5.4 Suitability of methods for measuring situation awareness

The goal of this research was to determine which method for measuring situation awareness is most suitable to be used for the evaluation of driver-vehicle interfaces in automated vehicles.

Contrary to our expectation the objective SAGAT-questionnaire showed no correlation with the amount of time available to respond to a situation. Moreover, it also did not seem to be correlated to the number of successful overtakes for each of the three time conditions. Contrary to the common assumption that SA improves as there is more time to regain it, the scores provided by the SAGAT questionnaire in our experiment showed a decline as available time increased. A possible explanation for this phenomenon is that the traffic in the lengthier time conditions was more changeable. This changeability may have resulted in confusion among the participants regarding which car was meant by the questions. A common confliction was between which car was driving in front and at the rear of the participant. Participants often reported the correct shape and colour of a car but in the wrong location. This may have been due to the fact that the time to orient on the vehicles was very short, and the confusion may have been a result of a too hasty observation. In any case it seems that the SAGAT questionnaire in the way it has been used in this experiment and in this setting is not suitable to measure SA.

The SART questionnaire has shown some promising results for use in the current set-up. Considering its advantages such as easiness of use and wide applicability, and based on the results of the research questions, we can conclude that the SART questionnaire appears to be the best indicator of SA for use in combination with automated vehicles.

6. Recommendations

From this research several suggestions for future research came forward which will be discussed in this chapter. Furthermore, an attempt was made to explain certain outcomes, which will also be explained here.

First, a possible explanation for the lack of correlation between the SA-scores provided by the SAGAT, and the criticality and successfulness of take-overs is that the results of the SAGAT are very susceptible to the moment at which the questions are asked. This moment may vary as participants had different amounts of time to respond to the situation itself. Furthermore it may also be of influence at which objects in the situation the questions are directed. It is possible that in the current task participants were generally only paying attention to a certain location which they believed to be of higher importance than another. For example a participant could be very focussed on the car in front as to not to collide with it, and meanwhile forget that he should pay attention to the car behind as well.

Secondly, the fact that the SA-scores provided by the SART were significantly lower than the scores provided by the SAGAT might be explained by the fact that people often seemed to get the feeling they were not paying enough attention if they were unable to respond to a SAGAT question. This may have affected them in the way they subjectively felt about the situation whilst answering the SART questionnaire, causing them to report lower levels of understanding than was actually the case.

A possible suggestion for future research is to investigate which of the measures is actually providing a false level of situation awareness. It might also be possible that both are accurate, but that the distribution of scores is more, or less, condensed in one of the measures. Another suggestion to increase the functionality of the SAGAT is to make the administration less confronting by letting participants give written answers, instead of verbal answers to the experimenter. This may reduce the feeling of a need to perform in front of the experimenter.

Concerning the SART questionnaire, an interesting consideration is that even though the SAGAT questionnaire did not correlate in any way with success-rate or criticality, it may have contributed to the validity of the SART-scores. Subjective measures, like the SART, are often scrutinised in literature. However, in the current experiment the SAGAT-questionnaire was administered first, which may have given the participants an indication of what they did or did not

notice in a certain situation. This may increase their insight into their own level of situation awareness and result in a more realistic self report. A suggestion for further research is to investigate a possible increase in validity of subjective measures when proceeded by a questionnaire implicitly providing an insight in the degree of understanding.

Furthermore, during the experiment it seemed there was a learning curve amongst the participants. Meaning they were at some point expectant of what was going to happen and were already ready to respond without there actually being a cause for such readiness. Although the current experiment did already take this in consideration by implementing several simulations in which no action was required, a suggestion might be to let the automatic system give a false alarm in several simulations. Through observation it should become clear whether the participant noticed the false alarm, or whether he automatically intervened without there being any reason to do so.

7. References

- Endsley, M. R. (1988). *Design and evaluation for situation awareness enhancement*. Paper presented at the Human Factors Society 32nd Annual Meeting, Santa Monica, CA.
- Endsley, M. R. (1988). *Situation Awareness Global Assessment Technique (SAGAT)*. Paper presented at the National Aerospace and Electronics Conference (NAECON), New York: IEEE.
- Endsley, M. R. (1989). A methodology for the objective measurement of situation awareness. Situational Awareness in Aerospace Operations (AGARD-CP-478) (Vol. 1, pp. 1 - 9). Neuilly Sur Seine, France: NATO - AGARD.
- Endsley, M. R. (1995). Direct Measurement of Situation Awareness in Simulations of Dynamic Systems. *Human Factors*, *37*(1), 65-84.
- Endsley, M. R. (1996). Automation and Situation Awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 163-181). Mahwah, NJ: Lawrence Erlbaum.
- Endsley, M. R., Kaber, D.B. (1999). Level of automation effects on performance, ,situation awareness and workload in a dynamic control task. *Ergonomics*, 42(3), 462-492.
- Endsley, M. R., Selcon, S.J., Hardiman, T.D., Croft, D.G. (1999). A comparative Analysis of SAGAT and SART for Evaluations of Situation Awareness.
- Hoeger, R. (2012). *Challenges in Automated Driving*. Paper presented at the Advanced Microsystems for Automotive Applications (AMAA), Berlin.
- Merat, N., Hamish Jamson, A. (2009). *How do drivers behave in a highly automated car*. Paper presented at the Driving Assessment 2009: 5th International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design Big Sky MT.

- Rauch, N., Kaussner, A., Krüger, H.-P., Boverie, S., Flemisch, F. (2009). *The importance of driver state assessment within highly automated vehicles*. Paper presented at the 16th ITS World Congress, Stockholm, Sweden.
- Saffarian, M., Winter, de, J.C.F., Happee, R. (2012, 56th Annual Meeting). *Automated Driving: Human-factors issues and design solutions*. Paper presented at the HUMAN FACTORS and ERGONOMICS SOCIETY.
- Salmon, P., Stanton, N., Walker, G., Green, D. (2006). Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics*, *37*, 225-238.
- Strybel, T. Z., Vu, K.L., Dwyer, J.P., Kraft, J., Ngo, T.K., Chambers, V., Garcia, F.P. (2007). *Predicting Perceived Situation Awareness of Low Altitude*
- Aircraft in Terminal Airspace Using Probe Questions. Paper presented at the Human Computer Interaction International Meeting, Beijing, China.
- Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design *Situational Awareness in Aerospace Operations (AGARD-CP-478)* (Vol. 3, pp. 1 - 17). Neuilly Sur Seine, France: NATO - AGARD.
- Verwey, W. B., Zaidel, D.M. (1999). Preventing drowsiness accidents by an alerness maintenance device. Accident Analysis and Prevention, 31, 199-211.
- Wiener, E. L., Curry, R.E. (1980). Flight-deck automation: Promises and Problems. *Ergonomics*, 23, 995-1011.

8. Appendixes

Index of Appendixes:

- 1. Form with demographic questions
- 2. Situation Awareness Rating Technique

Appendix 1

Ben jij een goede chauffeur?

Proefpersoonnummer:						
Datum:		Tijd:				
Geslacht:	M / V					
Leeftijd:						
In het bezit van een rijbewijs B sinds:						
Per maand rij ik gemiddeld o	ongeveer:					
KNA						
KIVI						
Omcirkel het kruisje/cijfer (dat het beste bij je past:					
Ik voel mij vandaag:						
Vermoeid XX	-XXXXXX	Uitgerust				
Ik ben een uitstekende chau	lffeur					
Oneens XX	-XXXXXXX	Eens				
Wanneer ik een auto bestuur kiik ik regelmatig in de spiegels						
Oneens XX	-XXXXXX	Eens				
Wanneer ik rij heb ik het gevoel dat ik de auto goed onder controle heb						
Oneens XX	-XXXXXX	Eens				
Wanneer ik rij heb ik over h	et algemeen een goed overzicht van	de verkeerssituatie				
Oneens XX	-XXXXXXX	Eens				

Geef aan

Hoe vaak ben je als bestuurder van een auto betrokken geweest bij een ongeluk?

.....

Hoe vaak was een ongeluk door jouw schuld of door gedeelde schuld?

.....

Geef met één of meerdere kruisjes aan waar volgens jou de 'dode hoek' zit bij een auto:



Ik geef mijzelf als chauffeur het cijfer:

1 2 3 4 5 6 7 8 9 10

Appendix 2

SART

1. Hoe **veeleisend** vond u de verkeerssituatie? (gelet op de complexiteit en veranderlijkheid van de situatie)

Helemaal niet

helemaal wel

2. In welke mate had u concentratie en/of inspanning nodig?

Weinig

veel

| X-----X-----X-----X-----X-----X

3. In welke mate heeft u tijdens de simulatie **inzicht verkregen** in de rijsituatie? (Weet u bijvoorbeeld welke auto's er in je directe nabijheid waren en wat die gingen doen?)

Weinig

veel