The Effect of Traffic on Situation Awareness and Mental Workload: Simulator-Based Study

Xueqin Hao[1,2], Zhiguo Wang[1,2], Fan Yang[1,2], Ying Wang[1,2], Yanru Guo[1,2], and Kan Zhang[1]

[1] State Key Laboratory of Brain and Cognitive Science, Institute of Psychology,
Chinese Academy of Sciences,
Beijing 100101, China
[2] Graduate School of the Chinese Academy of Sciences,
Beijing 100049, China
haoxq@psych.ac.cn

Abstract. In the present study, we investigated the effects of the different traffic on the driver's situation awareness and the mental workload (MWL). The task used in this study was a medium fidelity, 3-dimensional simulation of a driving environment. The simulation required participants to drive the user's car and perform a real-world driving task. After the simulated driving, participants were asked to complete two tests which assessed their situation awareness (SA). The mental workload measures in this study consisted of the physiological measures and the subjective assessment. Every participant performed two different traffic simulated driving conditions, one was low traffic, the other was high traffic. The results showed that with the increasing of traffic, the driving performance did not worsen, however participant's mental workload increased, at the same time, the participant's situation awareness performance deteriorated. Meanwhile, our results also demonstrated that recall-based SA test and recognition-based test was heterogeneous.

Keywords: situation awareness, workload, simulated driving.

1 Introduction

The research described here investigated situation awareness (SA) in the real-time task of driving. In particular, we are interested in how drivers monitor the location of vehicles around them. Situation awareness is an important concept in applied research, originating as a description what makes up the ace factor in 1980's [1]. With the development of complex machine system, it is difficult to entirely explain how the human error results in the disastrous accidents only according to those conventional cognitive concepts. Furthermore, the assessment of interface and system design is challenging in the complex system. Thus the cognitive concept of Situation Awareness (SA) offers a new perspective. Today, it is widely accepted that good SA is required for successful and safe performance in tasks such as flying and air traffic control (ATC) [2]. In contrast, low SA has been linked to operator error [3]. The

majority of situation awareness researches refer to flight and ATC. Although there are some similarities between the domains of flying and driving, there is paucity of empirical studies of SA in driving.

There are diverse definitions of situation awareness in the literature [4]. Arguably, the most accepted is by Endsley, who defines it as "The perception of 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 [5].

As Wickens (2001) suggests, the nature of "the situation" of which awareness is maintained needs to be specified [6]. Here we will deal only with the driver's awareness of the locations and properties of other nearby vehicles on the road way. Perhaps this should be called "traffic awareness".

Johannsdottir et al (2003) suggested a Cognitive Systems Engineering (CSE) framework and argued that the underlying representation must be distinguished from SA [7]. In the CSE framework the theoretical concept at the centre is the Dynamic Mental Workload (DMM) which captures the constant representation and updating of the system and the task environment in working memory (WM). The framework also includes three empirical or measured constructs that reflects on the DMM, namely SA, workload, and task-relevant performance. They suggested that using converging measures of the three central concepts: SA, workload, and task-relevant performance is more adequate for assessing the complex operator-machine interaction.

The researchers who focused on the traffic situation awareness seldom were concerned for the change of mental workload simultaneously. In the present study, participants' SA, workload, and driving performance were measured, the central objective here is that using converging measures to explore how the three constructs change with the change of the traffic.

There were two hypotheses in our study: with the increasing of the traffic, participants' mental workload become heavier, SA becomes worse, and driving performance deteriorated. It was also hypothesized that for the same driving scene, the SA performance firstly tested was better than one tested lastly.

2   Methods

2.1   Participants

Forty-three drivers (23 skilled drivers vs. 20 novices) participated in this experiment. The skilled drivers were those who had driven for more than 3 years with more than 20,000 km, while the novice drivers were those who just got their license with little driving experience. 22 were males, and 21 were females, aged from 20 to 37. The average was 28.6 years ($SD = 4.9$).

Latin square design was implemented to counterbalance potential order effect and learning effect.

2.2   Experimental Design

There were two different experimental designs for different dependant variables. For the four SA performance indexes and the driving performance index, a 2 (the order of tests: firstly situation judging test, firstly situation resuming test) × 2 (traffic: low,

high) × 2 (gender: male, female) × 2 (experience: novice, skilled) four-factor mixed experiment design was applied. For the workload indexes, a 2 (traffic: low, high) × 2 (gender: male, female) × 2 (experience: novice, skilled) three-factor mixed experiment design was used. While the driving experience and gender were between-subjects variables, and the order of test and the traffic were within-subjects variables.

In this study, participants were asked to perform two traffic conditions. In the low traffic, there were four vehicles on the screen around the user when the moving screen disappeared. In the high traffic, there were six vehicles on the screen when simulated driving task finished.

The order of tests was assigned as follows: In every traffic condition, participants needed to perform two blocks. In one block, participants were measured firstly with the situation resuming test (SRT), then with the situation judging test (SJT). In other block, the order of test was reversed. Every participant needed to finish 4 blocks experiments, in all 24 trials.

Dependent variables included: driving performance, SA performance, physiological indexes and subjective evaluations of MWL.

2.3 Materials

The experimental task was conducted using a medium fidelity, fully configured driving simulator, which was developed ourselves to measure SA performance and driving performance. It consisted of two parts. One part included a physical steering wheel, physical gas and brake pedals. Another part involved a PC-based simulated driving and SA measuring system.

The simulation environments included a three-lane highway presented from an egocentric viewpoint inside the driver's vehicle. Participants had full use of the rear view mirror as well as both side mirrors providing them with a complete view of the traffic behind them. Surrounding traffic was located on the two lanes driving in the same direction as the participant's car. On some trials, incident would occurred that required a driving performance , for example , a car could move into the driver's lane ahead of (or behind) while moving slowly (or fast) enough that it would collide with the user car. Participants could avoid hazards to accelerate or to decelerate, or move to the lane on the left, or move to the lane on the right. Each scene was scripted to stop at a random location after 60-80 seconds of run time. The scene was updated at a rate of 30 Hz. Drivers must keep track of their route location, the position and speed of their own and other vehicles. As participants were also asked to complete two SA measures which involved the awareness of position nearby vehicles. They were situation resuming test (SRT) which was similar with Gugerty's methods and situation judging test (SJT) which was developed ourselves [4]. At the same time simulator automatically recorded a series of data which consisted of the position and speed of user's car at every second. Furthermore, the simulator automatically collected data regarding the actual location of the cars relative to the participants' vehicle at the end of each trial.

SRT: In this test, the SA performance was assessed using a post-trial question (position grid) which the participants used to indicate the position of cars relative to their own vehicle at the time each trial ended on a bird's-eye view of the road. The bird's-eye view showed the road 20 car lengths ahead of or behind the driver. The

driver's car was in red in the correct lane. Participants could use mouse and Delete buttons to move or delete cars they had placed on the road. When participants finished recalling all the car locations for a scene, they clicked on the DONE button and then continued.

SJT: In this test, participants were presented four driving scenes and were asked to judge which scene was the correct car location scene at the time each trial ended. There was only one correct situation scene in the four driving scenes. Three driving scenes as distractors were produced by changing one car's location of correct driving scene. The changing rule was to move one car from its original location to the symmetrically opposite lane. Three cars moved were respectively in the near section, middle section, and far section. When participants finished four judgments, the next trial was automatically presented by the simulator.

Simulated driving task was presented with a Dell Latitude D500 notebook.

Physiological indexes included heart rate, heart rate variability index -LF/HF, which were measured by recording equipment -- KF2 dynamic multi-parameter bio-signal detector.

Subjective evaluations of MWL were measured by NASA-LTS questionnaire, including six items for six dimensions: mental demand, physical demand, time pressure, self-evaluation, effort and frustration. Participants were asked to give a value between 0 (not at all) and 20 (extremely) to indicate their subjective feelings in the temporal manipulation on every dimension.

2.4 Procedures

The participants were tested individually in a quiet room. Each participant completed the entire experiment in one day according to the following procedures: (1) 5 min of general instruction on the experimental task; (2) about 15 min of training, including three entire trials; (3) 5 min of instruction of the SA questionnaire and subjective workload rating scale to be administrated during experimental trials; and (4) four 15-min formal experimental blocks, including the two SA tests and the subjective workload rating. They were low traffic firstly SJT, low traffic firstly SRT, high traffic firstly SJT, and high traffic firstly SRT. The presenting order of these four blocks was balanced by Latin square to participants, however two conditions in same traffic was tested in succession.

The MWL subjective workload rating scale were administrated after every traffic conditions tasks, thus each participant had to take two times of evaluation on subjective MWL. The physiological recording began before the practice and stopped after the formal experiment.

3 Results

Here, the results of the traffic and the order of SA tests were presented. The results of individual difference about gender and driving experience would be presented in other article.

The SA data included four indexes: the score of resuming direction (SXY), the estimating error (RE), the total score of judging (SJ) and the react time of judging

(RT). Collision number (SB) was recorded automatically by simulator as driving performance.

Both x- and y-coordinates for each car were collected when the simulation was frozen at the end of each trial. The y-coordinates measured the distance between each car and the participants' car, while the x- coordinates measured whether the car was positioned in the right of or the light lane. These data were compared to the participants' evaluation using the grid positioning sheet.

For SXY, participants could get two points if they put a car into the correct quadrant regardless of the distance error. If the direction of x- or y- coordinate was right, they could get one point. Then sum up all car's scores of direction and averaged it then we got the SXY. RE were calculated as the difference between the computer value and the participants' estimation of car distance. Missed cars or added cars were therefore given the value of 120 meter (maximum error). SJ was the sum of four judging scores. Participant could get one point if one judgment was correct. The correct judgment only could get 0.5 point if participants gave two "true" judgments and one was correct. If participants gave "true" judgment more than two times, the correct judgment could not get points.

In the present study, we received 33 effective data at last (expert = 16, male = 8, female = 8; novice = 17, male = 9, female = 8). All trials from 33 participants, 24 trial each, 792 trials in total, were collected.

3.1 SA Performance

Table 1 showed the results of SJ performance in different conditions. The data reported in this article were analyzed by means of a repeated measures analysis based on the multivariate general linear model. A 2 (the order of tests: firstly SJT, firstly SRT) × 2 (traffic: low, high) × 2 (gender: male, female) × 2 (experience: novice, skilled) repeated measures analysis of SJ revealed significant main effects of traffic [$F(1,29) = 5.243$, $p < .05$] and gender [$F(1,29) = 6.398$, $p < .05$]. Whereas the main effect of order of tests [$F(1,29) = 2.942$, $p > .05$] was not significant. No interaction was significant. The results showed SJ decreased along with the increase of traffic.

Table 2 showed the results of RT performance in different conditions. The four factor repeated measures analysis of RT revealed significant main effects of traffic [$F(1,29) = 6.263$, $p < .05$], the order of tests [$F(1,29) = 11.934$, $p < .01$], and gender [$F(1,29) = 4.13$, $p < .05$], whereas all interactions were not significant. The results indicated RT increased along with the increase of traffic.

Table 1. The SJ in different conditions

SJ	N	Mean	SD
SJ_DP	33	3.30	0.49
SJ_DH	33	3.17	0.47
SJ_GP	33	3.09	0.54
SJ_GH	33	2.99	0.52

Table 2. The RT in different conditions

RT	N	Mean	SD
RT_DP	33	3.61	1.42
RT_DH	33	3.50	1.31
RT_GP	33	4.42	1.62
RT_GH	33	3.74	1.05

It is interesting that RT in firstly SJT condition was longer than the firstly SRT condition regardless of traffic. This was not consistent with our hypothesis.

Table 3 showed the results of SXY in different conditions. The four factor repeated measures analysis of SXY revealed significant main effects of traffic [$F(1,29) = 23.486$, $p < .001$], and the order of tests [$F(1,29) = 12.291$, $p < .01$], whereas all interaction wasn't significant. The results indicated that the more traffic resulted in the worse SXY performance. It is also interesting that SXY in firstly SRT condition was worse than the firstly SJT condition regardless of traffic. This was not also consistent with our hypothesis.

Table 3. The SXY in different conditions

SXY	N	Mean	SD
SXY_DP	33	0.86	0.48
SXY_DH	33	0.84	0.60
SXY_GP	33	0.78	0.82
SXY_GH	33	0.73	0.98

Table 4. The estimate error in different conditions

RE	N	Mean	SD
RE_DP	33	21.59	14.93
RE_DH	33	26.55	13.98
RE_GP	32	33.71	15.70
RE_GH	33	39.81	19.16

Table 4 showed the results of RE in different conditions. The four factor repeated measures analysis of RE revealed significant main effects of traffic [$F(1,29) = 47.426$, $p < .001$], and the order of tests [$F(1,29) = 21.082$, $p < .01$]. All interactions were not significant. The results indicated that RE increased along with the increasing of traffic. It is also interesting that RE in firstly SRT condition was larger than the firstly SJT condition regardless of traffic. This was not consistent with our hypothesis.

The results of the SA performance indicated that high traffic resulted in worse SA. There were also some unexpected result for the order of tests : Except for the RT

performance, the other three SA performances were better in conditions when participants implemented firstly SJT than those in conditions when participants implemented firstly SRT.

3.2 Driving Performance

Table 5 showed the results of SB in different conditions. The four factor repeated measures analysis of SB did not reveal any significant main effects whereas there was a significant interaction of EXPERICE * GENDER [$F(1,28) = 4.607$, $p < .05$]. The results showed that SB had not been influenced by the traffic and the order of tests.

 The driving performance showed that the traffic and the order of tests did not affect the driving performance in present study. Maybe this index was not enough sensitive for drivers' controlling performance.

Table 5. The collision number in different conditions

SB	N	Mean	SD
SB_DP	33	1.22	0.47
SB_DH	33	1.25	0.55
SB_GP	33	1.23	0.59
SB_GH	32	1.29	0.63

3.3 Workload Performance

Table 6 showed the results of six terms of NASA-TLX. In the workload data, we only analyses the effect of the traffic and gender and experience. The three factor(traffic, gender, and driving experience) repeated measures analysis of six dimension of NASA revealed significant main effects of traffic in three dimension which included mental required [$F(1,28) = 14.768$, $p < .001$], time pressure [$F(1,28) = 9.215$, $p < .001$], self-estimate [$F(1,28) = 6.009$, $p < .05$].   The results showed that with the increasing of traffic, the score of mental required, time pressure increased, whereas the score of self-estimate decreased.

Table 6. Subjective mental workload in different traffic conditions

Subjective mental workload	Low traffic		High traffic	
	M	SD	M	SD
MR	11.75	3.89	14.06	3.56
PR	6.156	4.0	6.69	3.6
TP	9.63	3.66	11.13	3.99
SF	9.81	3.76	8.16	3.9
EF	13.47	2.87	14.53	3.39
FR	8.97	3.98	10.25	4.04

3.4 Physiological Data

Table 7 showed the results of two physiological indicators. The three factor (traffic, gender, and driving experience) repeated measures analysis of HR revealed a significant main effects of traffic [$F(1,28) = 5.387$, $p < .05$]. The three factor repeated measures analysis of HRV revealed significant main effects of traffic [$F(1,28) = 4.872$, $p < .05$] and gender [$F(1,29) = 7.577$, $p < 0.01$]. The results showed that with the increasing of traffic, heart rate and heart rate variability increased.

Table 7. The physiological indexes in different traffic conditions

Physiological indexes	Low traffic		High traffic	
	M	SD	M	SD
Heart rate	72.22	8.28	73.43	9.81
LF/HF	2.41	1.41	2.77	1.55

4 Discussions

The results in present study indicated that with the traffic increasing, participants aroused more energy to cope with the complex task, while their subjective mental workload increased and their SA performance worsen. These results confirmed our hypothesis.

During stimulated driving, in order to track surrounding traffic, participants must shift attention to locations in the forward visual field as well as to rear- and side-view mirrors. Thus, this task requires not only perceptual tracking but also the use of dynamic spatial working memory. Therefore, when the traffic were increased, firstly drivers working memory load were added, secondly, the difficulty of avoiding hazardous also was added in that surrounding excessive vehicles diminished the safe field. In our study, two SA measures were memory-based. Thus, it is not surprising that the increased traffic resulted in worse SA performance.

But, there was a result which was not consistent with our hypothesis that participants' driving performance was not influenced significantly by the change of traffic. There were two reasons. Firstly, the reason is that we selected an index which it was not sensitive enough to the experiment manipulation. Thus, some more sensitive driving performance indexes would be necessary in the future, for example, departure distance. Secondly, the perceptual tracking subtask was not same affected by the memory workload as the memory-based SA performance.

SA performance which was measured with two methods indicated SA declined when the traffic increased. This implied that SA is a sensitive index. Furthermore, these results were consistent with Gugerty's results. But, there were some difference between the two studies. Generally, we decreased participants' memory load in SA tests. In SRT, participants were supplied the amount and color of surrounding cars, they only need put these cars on the correct locations. In SJT, participants were presented four choices, they only need to judge which one was correct. In other words, SJT was a recognition task.

We hypothesized that the SA performance of measure firstly implemented would better than those of tested lastly, but it was not the case. The results in present study were inconsistent with our hypothesis: Except for the RT in SJT, other three SA performance were better in conditions which participants were measured firstly with SJT than those conditions which participants were measured firstly with SRT. Maybe it result from the different dependence of two tests on participants' memory: In SJT, participants were presented four driving scenes, driving scenes presented in SJT were radically consistent with participants' mental model about driving task. In other words, the driving scenes would be helpful to maintain participants' mental model. While the resuming test did not provide a replaceable perception represent. So SJT would slow the vanishing of the participants' driving dynamic mental model. Thus our results also indicated that two SA tests were heterogeneous.

References

1. Kamilla, B.A.: Situation Awareness and Working Memory: an Integration of an Applied Concept with a Fundamental Cognitive Process, dissertation (2004)
2. Endsley, M.R.: Theoretical Underpinnings of Situation Awareness Analysis and Measurement. Lawrence Erlbaum Associates, Mahwah, NJ (2000)
3. Jones, D.G., Endsley, M.R.: Sources of situation awareness errors in aviation. Aviation, Space and Environmental Medicine 67, 507–512 (1996)
4. Mathews, M.L., Byrant, D.J., Webb, R.D.G.: Situation Awareness and Driving?Literature Review and Analysis (No. T8080-8-1348/001/ss). Transport Canada, Montreal, QB (1999)
5. Endsley, M.R.: Measurement of situation awareness in dynamic systems. Human Factors 37, 65–84 (1995)
6. Wickens, C.D., Hollands, J.G.: Engineering Psychology and Human Performance. Prentice Hall, New Jersey (1999)
7. Johannsdottir, K.R., LeEevre, J.-A., Herdman, C.: Cognitive system engineering (CSE) framework for evaluating cockpit interfaces. In: The Proceedings of The Twelfth International Symposium on Aviation Psychology, Colombus, OH, The Ohio State University (2003)
8. Gugerty, L.J.: Situation awareness during driving: Explicit and implicit knowledge in dynamic spatial memory. Journal of Experimental Psychology Applied 3, 42–46 (1997)