Presented at the HCI International 2005 (1^{st} International Conference on Augmented Cognition)
August 25-27, 2005.

Human Interaction with Adaptive Automation: Strategies for Trading of Control under Possibility of Over-trust and Complacency

Toshiyuki Inagaki, Hiroshi Furukawa, Makoto Itoh

Department of Risk Engineering, University of Tsukuba
Tsukuba 305-8573 Japan
{inagaki, furukawa, itoh}@risk.tsukuba.ac.jp

Abstract

Function allocation needs to be dynamic and situation-adaptive to support humans appropriately. Machines have thus been given various types of intelligence. It is true, however, humans working with such smart machines often suffer negative consequences of automations. By taking the adaptive cruise control (ACC) system and the lane keeping support (LKS) system as examples of adaptive systems in the real world, this paper investigates human trust in and reliance on adaptive automation, and discusses design of strategies for trading of control under possibility of driver's over-trust in automation.

1 Introduction

The design decision of assigning functions to human and machine is called *function allocation*. In spite of its importance, function allocation is still a kind of art. Various strategies for function allocation have been proposed. Traditional function allocation strategies are classified into three types: (a) comparison allocation, (b) leftover allocation, and (c) economic allocation. These strategies consider "who does what." Such design decisions yield function allocations that are *static*: viz., once a function is allocated to an agent, the agent is responsible for the function at all times. However, the operating environment may change as time goes by, or performance of the human may degrade gradually as a result of psychological or physiological reasons, which suggests that "who does what" design decisions are not sufficient and that "who does what and when" considerations are needed. If the total performance or safety has to be maintained rigorously, it may be wise to reallocate functions between the human and the machine: viz., the resulting function allocation must be dynamic.

A scheme that modifies function allocation dynamically depending on situations is called an *adaptive function allocation*. The adaptive function allocation assumes criteria to determine whether functions have to be reallocated, how, and when. The criteria reflect various factors, such as changes in the operating environment, loads or demands to operators, and performance of operators. The automation that operates under an adaptive function allocation is called *adaptive automation* (Inagaki, 2003; Moray, Inagaki, & Itoh, 2000; Parasuraman, Bhari, Deaton, Morrison, & Barnes, 1992; Rouse, 1988; Scallen & Hancock, 2001; Scerbo, 1996). Adaptive automation works intelligently and reliably. It can sense and analyze situations, decide what must be done, and implement control actions. It may be easy for humans to trust in such a smart machine. However, if human-computer interaction strategies were poorly designed, various types of automation surprises may occur. Based on our experiments with the adaptive cruise control (ACC) system and the lane keeping support (LKS) system, this paper investigates human trust in and reliance on adaptive automation, and discusses design of strategies for trading of control under possibility of driver's over-trust in automation.

2 ACC and LKS Systems in Advanced Automobile

The adaptive cruise control (ACC) system and the lane keeping support (LKS) system may be regarded as real-world examples of adaptive systems. The ACC system is a partial automation for longitudinal control, designed to reduce the driver's workload by freeing the driver from frequent acceleration and deceleration. It controls the *host vehicle* so that it can follow a vehicle ahead (*the target vehicle*) at a driver-specified distance by controlling the engine and/or power train and potentially the brake. When the ACC system detects the deceleration of the target vehicle, it slows down the host vehicle at some deceleration rate. As long as the deceleration of the target vehicle stays within a certain range (say, not greater than 0.2G), the ACC system can control the speed of the host vehicle perfectly and no rear-end collision into the target vehicle occurs. However, when the target vehicle makes a rapid deceleration at a high rate (e.g., 0.4G), the ordinary brake by the ACC system may not powerful enough to avoid a collision into the target vehicle. How to prepare for such cases is one of significant design issues.

The LKS system is a partial automation for lateral control, designed to reduce the driver's workload to keep the host vehicle within its driving lane. Recognizing the lateral position of the host vehicle in the lane, the LKS system assists the driver's steering control to keep the car center of the lane with the use of the power steering system.

3 Two Key Concepts Needed for Systematic Investigations

This section gives a brief summary of two key concepts, *trading of control* and *levels of automation*, that are needed for investigating human interactions with adaptive systems in a systematic manner.

3.1 Trading of Control

Trading of control (or, *trading of authority*) refers to the human-computer collaboration in which either one of the human or the computer is responsible for a function, and an active agent changes alternately from time to time. Trading of control is essential for adaptive automation, because adaptive automation needs to modify function allocation between humans and machines dynamically in response to changes in situations, human workload, or performance. A scheme to implement trading of control is called an *automation invocation strategy*. There are some types of automation invocation strategies (Inagaki, 2003). Among them, this paper discusses the following two classes that are important for transportation systems.

Critical-event strategies: Automation invocation strategies of this class change function allocations when specific events (called, *critical events*) occur in the human-machine system. It is assumed that human workload may become unacceptably high when the critical events occur. If the critical events did not occur during the system operation, allocation of functions would not be altered.

Measurement-based strategies: Automation invocation strategies of this class adjust function allocation dynamically by evaluating moment-to-moment workload or total system performance. It is necessary to develop custom tailored algorithms if the system is to be compatible with individual operators. Individual differences in human operator capabilities will also influence the response to multiple task demands.

3.2 Levels of Automation

In describing various types of human-computer interactions, the notion of the *level of automation* (LOA) is useful. Table 1 gives an expanded version in which a new LOA comes between levels 6 and 7 in the original list by Sheridan (1992). The added level, called the level 6.5, has been firstly introduced in (Inagaki, Itoh, & Moray, 1997) with two-fold objectives: (1) to avoid automation surprises that may be induced by automatic actions and (2) to implement actions that are indispensable to assure systems safety in emergency.

Table 1: Scales of Levels of Automation (expanded version)

1. The computer offers no assistance; human must do it all.
2. The computer offers a complete set of action alternatives, and
3. narrows the selection down to a few, or
4. suggests one, and
5. executes that suggestion if the human approves, or
6. allows the human a restricted time to veto before automatic execution, or
6.5 executes automatically upon telling the human what it is going to do, or
7. executes automatically, then necessarily informs humans,
8. informs him after execution only if he asks,
9. informs him after execution if it, the computer, decides to.
10. The computer decides everything and acts autonomously, ignoring the human.

4 Four Dimensions of Trust

Lee & Moray (1992) distinguished four dimensions of *trust*: (1) *foundation* that represents the “fundamental assumption of natural and social order that makes the other levels of trust possible,” (2) *performance* that rests on the “expectation of consistent, stable, and desirable performance or behavior,” (3) *process* that depends on “an understanding of the underlying qualities or characteristics that govern behavior,” and (4) *purpose* that rests on the “underlying motives or intents.”

It would not be hard for a human operator to recognize that the designer’s purpose or intention in introducing adaptive automation lies in regulating operator workload at some optimal level. Respecting the second and the third dimensions may not be straightforward: Human’s understanding of the automation invocation algorithm may be imperfect if the algorithm is sophisticated or complicated. When the human failed to be certain of the second and the third dimension of trust, she/he would fail to establish trust in the adaptive automation. Human’s distrust or mistrust in automation can cause inappropriate use of automation, as has been pointed out by Parasuraman and Riley (1996).

5 Trust in and Reliance on Automation: Experiment 1

This paper discusses human trust in adaptive systems from the viewpoint of *trading of control* and the *levels of automation*. The first two experiments deals with human-automation interactions and trading of control in emergency, and the third experiment discusses those in peacetime. Now, here comes Experiment 1.

5.1 Strategy for Trading of Control in Emergency

Consider a low-speed range ACC system with automatic stopping functionality that controls the host vehicle so that it can follow a target vehicle at a pre-specified distance in a heavy or jammed traffic. The specified distance can be maintained successfully as long as the acceleration/deceleration rate of the target vehicle stays within a certain range. When the target vehicle makes a rapid deceleration to which the ACC’s ordinary brake is not strong enough to avoid a crash, and if the ACC system was the only agent to cope with the situation, a rear-end collision may occur eventually. Let us consider the following scheme for trading control from the ACC system to the driver.

Scheme 1: Upon detecting a rapid deceleration of the target vehicle, the ACC gives an *emergency-braking alert* that tells the driver to hit the brake her/himself hard enough to avoid a collision.

5.2 Participants and Performance Measures

The purpose of Experiment 1 is to investigate drivers’ trust in and reliance on the emergency-braking alert. Twenty students participated in the experiment. The subjects were divided into two groups: (1) Group 1 of ten people and (2) Group 2 of ten. Each subject received 17 trials, each of which lasted about 3 minutes. During Trials 1 to 3, no emergency-braking alert functionality was available: Subjects had to judge themselves whether she/he has to take over control from the ACC system and when. During Trials 4 to 10, a correct emergency-braking alert was issued immediately when the target vehicle decelerated rapidly. In Trial 11, an emergency-braking alert was missed when the target vehicle made a rapid deceleration. In Trial 12, subjects in Group 1 received a correct emergency-braking alert, while alert was missed again for subjects in Group 2 upon the target vehicle’s rapid deceleration. During Trials 13 to 17, correct alerts were given to all the subjects in either group.

The following data were recorded as performance measures. (1) Response time; viz., time elapsed before the subject initiated to apply the brake, (2) number of unnecessary brakes applied when no rapid deceleration was made by the target vehicle, (3) subjective ratings of trust in emergency-braking alerts, (4) subjective ratings of usefulness of emergency-braking alerts, and (5) what subjects said during and between trials.

5.3 Results

The experiment was conducted with a fixed-based driving simulator. Mean response times of the subjects of the two groups are shown in Figure 1.

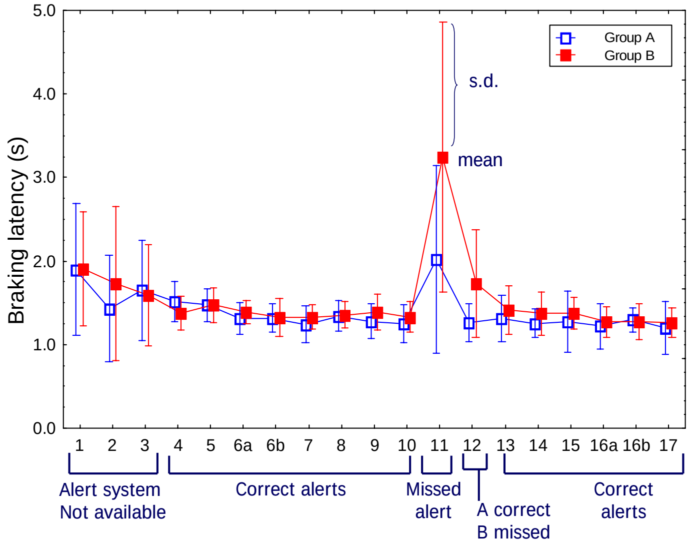
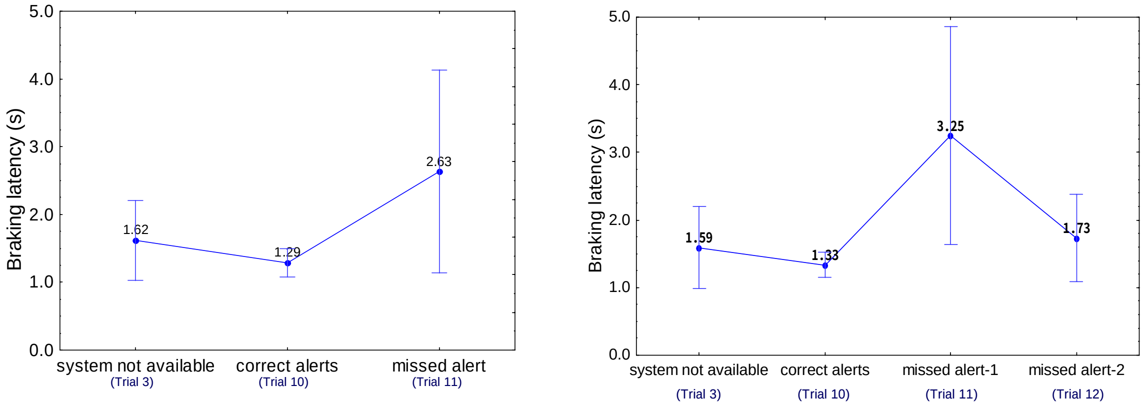


Figure 1: Response Times for Trials 1 through 17

(1) Significant differences were found among the mean response times of 20 subjects for the cases of Trials 3, 10, and 11 (F(2, 57) = 11.19, p < .0001); see, Figure 2(a). While receiving correct emergency-braking alerts consecutively and repetitively, subjects became reliant on the alerts, and tended to judge based on the alerts whether an emergency brake was necessary or not. When a correct alert was missed at Trial 11, subjects showed significantly late responses to the rapid deceleration of the target vehicle.

(2) Subjects in Group 2 experienced two consecutive missed alerts at Trials 11 and 12. Significant differences were found among response times of subjects for Trials 3, 10, 11, and 12 (F(3, 36) = 8.86, p < .0002); see, Figure 2(b). Among those response times, the response time for Trial 11 differed significantly from others, and no significant differences between response time for Trial 12 and either one of those for Trials 3 and 10, which implies that, when subject experienced a missed alert, they became vigilant for a while and tried to judge the situation to determine whether an emergency brake was in need.



(a) All Subjects

(b) Group B

Figure 2: Effects of Missed Alert(s) on Response Time

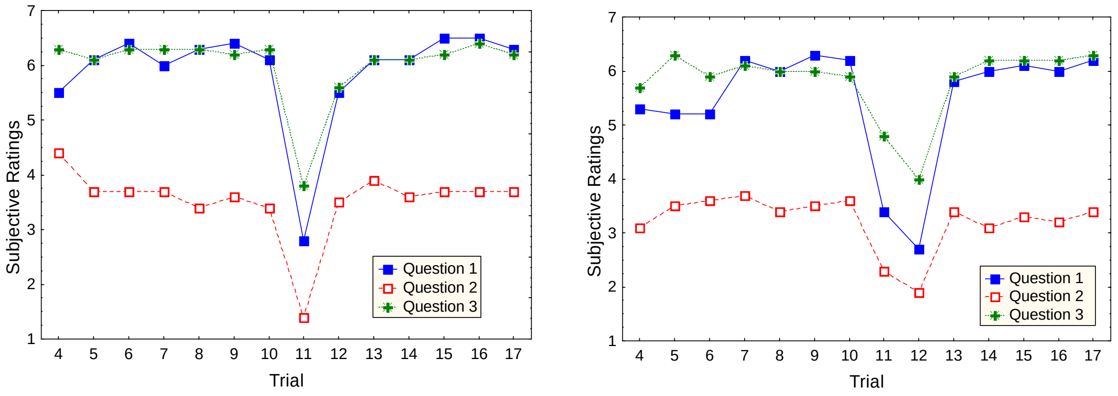
(3) A McNemar test found a significant difference between the numbers of unnecessary brakes before and after the missed alert at Trial 11, which suggests that subject behaviors differ between before and after experiencing the missed alert. The same situations were placed symmetrically around Trial 11: viz., each pair of trials (8, 14), (9, 13), and (10, 12) are essentially the same.

(4) Protocol data indicate that subjects' policies in use of alerts changed after experiencing the missed alert. Among 27 protocol data collected before the missed alert trial, 15 protocol data tell that many subjects used emergency-braking alerts to evaluate the situation. After experiencing the missed alert, the number of alert-reliant protocol data was reduced to just 6 out of 28. On the other hand, 20 out of 28 protocol data suggest that subjects tried to evaluate circumstances themselves and used alerts as supplements or confirmation to their judgements.

(5) As described in section 4, *trust* has four dimensions. The authors believe that these dimensions should be distinguished in the subjective rating of trust. Since the first and the fourth dimensions are trivial for emergency-braking alerts, participants of Experiment 1 were requested to evaluate the second and the third dimensions of trust in 7-point scales. They were asked: (Q1): "To what extent do you figure out why an emergency-braking alert was issued (or why it was missed) when the target vehicle made a rapid deceleration?" (1: not at all, 4: partly, 7: completely). (Q2): "Were the alerts given at the right time?" (1: too late, 4: exactly when I expected, 7: too early). Q1 deals with the third dimension of trust (i.e., *process* that depends on "an understanding of the underlying qualities or characteristics that govern behavior"), and Q2 the second dimension (i.e., *performance* that rests on the "expectation of consistent, stable, and desirable performance or behavior"). As for the subjective ratings of usefulness of emergency-braking alerts, the participants were asked: (Q3): "To what extent do you think is this warning system useful?" and requested to answer in the 7-point scale (1: not at all, 4: not sure, 7: very useful).

Figure 3(a) depicts means of the ratings of subjects in Groups A, and Figure 3(b) those of subjects in Group B. All the ratings for Q1 through Q3 behave in a similar manner. However, the rated values for Q2 are smaller than those for Q1 or Q3, which suggests that the subjects were not satisfied with the timings of the emergency-braking alerts.

It is also observed in Figures 3(a) and 3(b) that the subjective ratings of trust in the automated alert recovered quickly, which suggests that it may not be wise to assume that drivers shall become vigilant and continues to be so once they experienced a missed alert. The effects of missed alerts do not remain long and subjects may start relying on the automated alerts again. This implies the need for technological backup measures for ensuring safety of the automobile.



(a) Group A (b) Group B

Figure 3: Subjective Ratings

6 Trading of Control under Driver's Over-trust in Automation: Experiment 2

The second experiment via a computer simulation approach investigated efficacy of strategies for trading of control in emergency under possibility of driver's over-trust in automation. The experiment tries to investigate the need for technological backup measures for ensuring automobile safety.

6.1 Driver's Psychological States and Their Transitions

When the target vehicle slows down at a certain deceleration rate not greater than 0.2G, the ACC system of the host vehicle begins to decelerate 0.5s after the initiation of the target vehicle's deceleration, in which the ACC system applies the brake at the rate of 0.2G. It would be natural for the driver to trust in the ACC system if she/he observed that the ACC system behaves correctly and appropriately. Sometimes the driver may place excessive trust in the automation. In such cases, the driver may fail to allocate her/his own attention to the driving environment, and may pay attention inappropriately to some non-driving tasks (e.g., use of a mobile phone, manipulation of on-board audio systems). Five psychological states are distinguished for the driver based on a model by (Hashimoto, 1984) as shown in Figure 4.

State IV: Hyper-normal and excited state.

State III: Normal and vigilant state
(the best state for safe driving).

State II+: Normal and relaxed state with
moderate level of trust in the
automation.

State II: Normal and relaxed state, with
complete faith in the automation.

State I: Subnormal and inactive state.

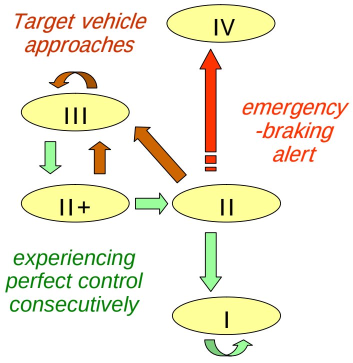


Figure 4: Driver's Psychological States and Their Transitions

It is assumed that, when the driver starts driving, her/his psychological state is positioned initially at State III (the best for safe driving), and that the state changes dynamically as time passes by; see, Figure 5. The following two types of events are distinguished for cases in which the ACC system can perform its longitudinal control with the ordinary brake: (1) *Type-1 event*, in which the target vehicle decelerates at a certain rate and the ACC system performs its longitudinal control perfectly so that it may satisfy the driver completely, and (2) *Type-2 event*, in which the target vehicle decelerates at a rate within 0.2G. Though the ACC system can cope with the situation within the ordinary brake, the driver feels alarm because the host car comes close to the target vehicle.

The contributions of the events to the driver's psychological state transitions are defined as follows: If the driver experiences a certain number (say, 10 or 20) of Type-1 events consecutively, the driver's psychological state changes one step downward (e.g., from III to II+). When the driver experiences a Type-2 event, her/his psychological state changes one step upward (e.g., from II+ to III). State I is the absorbing state: Once the driver enters into State I, she/he never comes out of the state even if a Type-2 event occurs.

6.2 Designs Alternatives for Trading of Control in Emergency

Suppose the ACC system recognizes that the deceleration rate of the target vehicle is much greater than 0.2G, the maximum deceleration rate to which the ACC system can cope with the ordinary automatic brake. In addition to Scheme 1 that was discussed in section 3, the following design alternatives may be feasible:

Scheme 2: Upon recognition of a rapid deceleration of the target vehicle, the ACC system gives an emergency-braking alert, and if the driver does not respond within a pre-specified time (2s in this paper), it applies an automatic emergency brake that decelerates at the rate of 0.4G.

Scheme 3: Upon recognition of a rapid deceleration of the target vehicle, the ACC system applies its automatic emergency brake to implement the deceleration rate of 0.4G simultaneously when it issues an emergency-braking alert.

Scheme 4: Upon recognition of a rapid deceleration of the target vehicle, the ACC system firstly applies its automatic emergency brake with the deceleration rate of 0.4G. Then, the ACC system tells the driver that it applied an emergency brake some seconds ago.

In term of the levels of automation, it can be said that the LOA = 4 for Scheme 1, LOA = 6 for Scheme 2, LOA = 6.5 for Scheme 3, and LOA = 7 for Scheme 4. Note that, from the viewpoint of swiftness of the emergency brake, Schemes 3 and 4 are indifferent. However, in case of Scheme 4, the driver may fail to recognize what is going on, when the ACC system applies its emergency brake. Scheme 3 is designed to avoid delay in the automatic emergency brake as well as to avoid automation surprises. Scheme 4 thus shall not be investigated further in this paper.

6.3 Model for Driver's Response to the Alert

When the driver hits the brake hard enough, she/he can make a deceleration at the rate of 0.5G. The driver's response time to the emergency-braking alert varies depending on the psychological state at that time moment:

(1) If the driver was in State I when the alert was set off, he/she does not respond to the alert at all. (2) If the driver was in State II, she/he stays in the state with probability 0.8, and hits the brake pedal in T2 seconds. With probability 0.2, the driver state changes to State IV. (3) If the driver was either in State II+ or in III, she/he applies the emergency brake her/himself either in T2+ or in T3 seconds, respectively. (4) In State IV, the driver panics and fails to take any meaningful actions to attain car safety.

T2, T2+, and T3 are random variables with different means. In the Monte Carlo simulations described in the following section, we assumed: T2 is uniformly distributed over the interval [2.7s, 3.3s], with the mean 3.0s, T2+ is uniformly distributed over [1.8s, 2.2s], with the mean 2.0s, and T3 is uniformly distributed over [1.35s, 1.65s], with the mean 1.5s.

6.4 Monte Carlo Simulations

The experiment has a 2 x 3 x 3 x 2 factorial design, mapping onto (*Headway Distance*) x (*Level of Automation*) x (*Event-Mixture Ratio*) x (*Driver's Psychological State Transition Condition*). Two levels, 80m and 50m, were distinguished for the *Headway Distance* between the host and the target vehicles. Three *Levels of Automation*, LOA-4, LOA-6, and LOA-6.5, were distinguished as the design alternatives for cases of a rapid deceleration of the target vehicle, corresponding to Schemes 1 through 3, respectively. The *Event-Mixture Ratio* denotes the proportion of the number of Type-2 events to the total sum of the numbers of Type-1 and type-2 events before the target vehicle makes a rapid deceleration. The Event-Mixture Ratio was set at 0%, 10%, or 20%. Two cases, 10 and 20, were investigated for the *Driver's Psychological State Transition Condition*: In cases of the former, if the driver experiences 10 Type-1 events consecutively, her/his psychological state changes one step downward. In the latter, 20 consecutive Type-1 events make a downward state transition.

The *Trip Length* was measured in terms of the number of the total sum of the Type-1 and Type-2 events before the target vehicle makes a rapid deceleration. In the present study, the Trip Length was fixed at 100; see Figure 5.

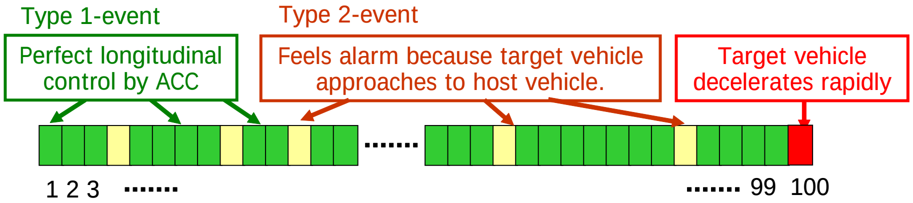


Figure 5: Events in a Trip

For each combination of the above conditions, 5000 Monte Carlo runs were performed with a discrete-event simulation software, WinCrew (Micro Analysis and Design, Inc.), under the following assumptions: The target vehicle makes a rapid deceleration at the rate of 0.4G at some time point while it has been running at the speed of 100km/hr. At the time of the rapid deceleration of the target vehicle, the host vehicle has also been running at 100km/hr, following the target vehicle. When detecting a deceleration of the target vehicle, the ACC system firstly applies its ordinary brake. One second after initiation of the ordinary brake, the ACC system recognizes that the deceleration of the target vehicle cannot be handled by the ordinary brake, and performs safety control action by applying one of Schemes 1 through 3.

6.5 Simulation Results

Table 2 shows the number of accidents (collisions into the target vehicle) during 5000 Monte Carlo runs and the distribution of the driver's psychological states when the trip length was 100.

Table 2: Computational results of automobile safety at the rapid deceleration of the target vehicle

LOA	Psychological state transition condition	Event-mixture ratio	# Runs ended in collision		Driver's psychological state at the emergency-braking alert				
			Headway distance: 80m	Headway distance: 50m	State I	State II	State II+	State III	State IV
4	10	0%	5000	5000	5000	0	0	0	0
		10%	1300	1478	1232	264	1003	2433	68
		20%	28	41	24	19	415	4538	4
	20	0%	4203	4786	3999	797	0	0	204
		10%	9	19	1	16	455	4520	8
		20%	0	0	0	0	34	4966	
6	10	0%	0	5000	5000	0	0	0	0
		10%	0	1280	1218	282	987	2451	62
		20%	0	36	27	32	415	4517	9
	20	0%	0	4192	3981	808	0	0	211
		10%	0	7	3	18	432	4543	4
		20%	0	0	0	0	31	4969	0
6.5	10	0	0	0	5000	0	0	0	0
		10%	0	0	1307	259	984	2375	75
		20%	0	0	22	27	426	4519	6

Let us take a look at the results when the headway distance was 80m. Comparing LOA-4, LOA-6, and LOA-6.5 for the cases where the Driver's Psychological State Transition Condition was set at 10, the number of accidents under LOA-4 was siginificantly larger than either one of those under LOA-6 and LOA-6.5, especially when the Event-Mixture Ratio is small. The result may be interpreted as follows: When the driving is peaceful in which Type-2 events seldom occur and the ACC system continues to be successful in its longitudinal control, the driver is likely to trust in and rely on the ACC system and his/her vigilance may degrade (see, Table 2 for the Driver's Psychological State at the time of the rapid deceleration). If the target vehicle makes a rapid deceleration and if an emergency-braking alert is given when the driver is in such a state with degraded vigilance, the driver may fail to cope with the circumstance in a timely manner.

What happened, when the headway distance was reduced to 50m? It is seen in Table 2 that, in case of LOA-6.5, the number of accidents remains unchanged (naught), irrespective of values of the Event-Mixture Ratio. However, in case of LOA-6, accidents occurred. Moreover, the increase in the number of accidents under LOA-6 is drastic when the headway distance is reduced from 80m to 50m. Why was LOA-6 not effective? The reason lies in the *time-delay* introduced for enabling the driver to initiate her/his emergency brake in that time period (e.g., 2s in the current simulation study). Note here that a discussion such as, "Is it better to make the time-delay shorter (say, to 1s)?" may not always be meaningful. For instance, ask yourself whether 1s is appropriate and long enough for the human to catch an alert, understand its implication, and implement a proper safety control action.

From the viewpoint of trust in and reliance on automation, an interesting observation may be made by comparing the number of accidents for the following two settings: (1) LOA-6 where the Driver's Psychological State Transition Condition was set at 10 and (2) LOA-4 where the Driver's Psychological State Transition Condition was set at 20. In the latter case, the number of accidents was 19, while 1280 accidents occurred in the former case. When the driver is responsible in applying the emergency brake at all times and on every occation, the driver may stay vigilant than in cases in which the driver can expect that the automation takes safety control actions when necessary. If this is the case, automobile safety may be degraded by a design decision to adopt a high LOA for safety control in emergency situations. However, that does not necesarily mean that theLOA has to be kept low. There are two pieces of evidence: First, even under LOA-6, the number of accidents can be reduced, if the Driver's Psychological State Transition Condition may be improved from 10 to 20 by some means, such as by adopting a better design of human interface: The number of accidents was reduced drastically from 1280 to 7 (see, Table 2). Second, the difference between the number of accidents (0) under LOA-6.5 and the number of accidents (19) under LOA-4 is statistically significant.

7 Automation Surprises in the Real World: Experiment 3

With various cars equipped with ACC and LKS systems, the authors conducted Experiment 3 to investigate human interactions with adaptive automation in peacetime. During the experiment, the authors did not read the owner's manuals thouroughly, which is usually the case for an ordinary car driver. While driving by letting the ACC and the LKS systems at work on expressways in Metropolitan Tokyo area, one of the authors (Inagaki) *simulated* various types of drivers, such as active/relaxed/lazy drivers. One day, when he was simulating a highly strained driver who holds the steering wheel tightly, he heard a voice message of the computer on the host vehicle. However, he failed to catch the exact meaning of it. Some time later, he saw a mild curve ahead, and expected the LKS system would steer the wheel appropriately. However, the author's trust was inappropriate: He was surprised to see the host vehicle went straight, passing across the lane boundary. By that time point, the computer had already traded the authority of lateral control from the LKS system to the driver. The computer had determined, by monitoring moment-to-moment steering torque of the driver, that he was not active in driving, and had decided it appropriate to return the steering task to the driver, according to a measurement-based strategy for trading of control in adaptive automation. The problem in this case lies in the fact that the trading of control was executed at a high LOA, by *assuming* (without confirming) that the driver should accept the trading and understand the need of it.

How can the computer communicate with the driver in such a circumstance? Several design alternatives with different LOA are possible: (1) The computer tells the driver, "You seem to be bored," where the LOA is set at 4. (2) The computer says, "Shall I let you drive yourself?" where the LOA is set at 5. (3) The computer gives a stronger message, "I will hand over control to you in a few seconds," where the LOA is positioned at 6. (4) The computer tells the driver, "I have just handed over control to you," where the LOA is set at 7. (5) The computer trades control from the LKS system to the driver silently. The LOA of this design is set at 8 or higher. It was the design alternative (3) that the author experienced in Experiment 3.

8 Concluding Remarks

This paper has discussed some issues of human trust in and reliance on adaptive automation by taking the ACC and the LKS systems as examples. Adaptive automation offers flexible human-machine interactions in a situation-dependent and context-specific manner. This very flexibility, however, may bring various inconveniences or undesired effects. This paper has shown that human-automation interaction design and its effects on humans can be investigated in a unified manner with two key concepts, *trading of control* and *levels of automation*. As the above examples in Experiments 1 through 3 illustrate, *value* or efficacy of LOA varies depending on situational context, such as time-criticality in the given circumstance. In other words, alothough trading of control at a high LOA may be needed in emergency, trading of authority at such a high LOA in peacetime may lead to possible loss of situation awareness or automation surprises. If the LOA was chosen inappropriately, some undesirable result may come out. In designing human-machine systems, it is important to predict how the design may affect humans and change their behaviors. The following three approaches are available for selecting an appropriate LOA: (a) cognitive experiment, (b) theoretical analysis, and (c) computer simulation. For details with illustrative examples, refer to (Inagaki, 2005).For further discussions on the design of human-automation interactions, refer to (Inagaki & Stahre, 2004).

Since 2004 July, the first author has been conducting a 3-year research project, entitled "Situation and intention recognition for risk finding and avoidance: Human-centered technology for transportation safety," with the research budget (5.3 Million US Dollars) supported by the MEXT of Japanese Government. One of the aims is to develop adaptive automation for automobile with critical-event and measurement-based automation invocation strategies with the aid of real-time technologies for sensing traffic environment as well as driver's psychological states, such as inappropriate trust in automation (see, Figure 6). Research results shall be reported in the near future.

References

Inagaki, T. (2003). Adaptive automation: Sharing and Trading of Control. In E. Hollnagel (Ed.) *Handbook of Cognitive Task Design* (pp. 147-169). LEA.

Inagaki, T. (2005). Design of human-machine interactions for enhancing comfort and safety. *Proceedings of the AAET*, 23-39.

Inagaki, T., & Stahre, J. (2004). Human supervision and control in engineering and music: Similarities, dissimilarities, and their implications. *Proceedings of the IEEE*, 92(4), 589-600.

Inagaki, T., Moray, N., & Itoh, M. (1997). Trust and time-criticality: Their effects on the situation-adaptive autonomy. *Proceedings of the AIR-IHAS Symposium*, 93-103.

Lee, J.D. & Moray, N. (1992). Trust, control strategies and allocation of function in human machine systems. *Ergonomics*, 35(10), 1243-1270.

Moray, N., Inagaki, T., & Itoh, M. (2000). Adaptive automation, trust, and self-confidence in fault management of time-critical tasks. *Journal of Experimental Psychology: Applied*, 6(1), 44-58.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230-253.

Parasuraman, R., Bhari, T., Deaton, J.E., Morrison, J.G., & Barnes, M. (1992). *Theory and design of adaptive automation in aviation systems* (NAWCADWAR-92033-60). Naval Air Development Center Aircraft Division.

Rouse, W.B. (1988). Adaptive aiding for human/computer control. *Human Factors*, 30(4), 431-443.

Scallen, S.F. & Hancock, P.A.(2001). Implementing adaptive function allocation. *International Journal of Aviation Psychology*, 11(2), 197-221.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.). *Automation and human performance* (pp.37-63). LEA.

Sheridan, T. B. (1992). *Telerobotics, automation, and human supervisory control*. MIT Press.

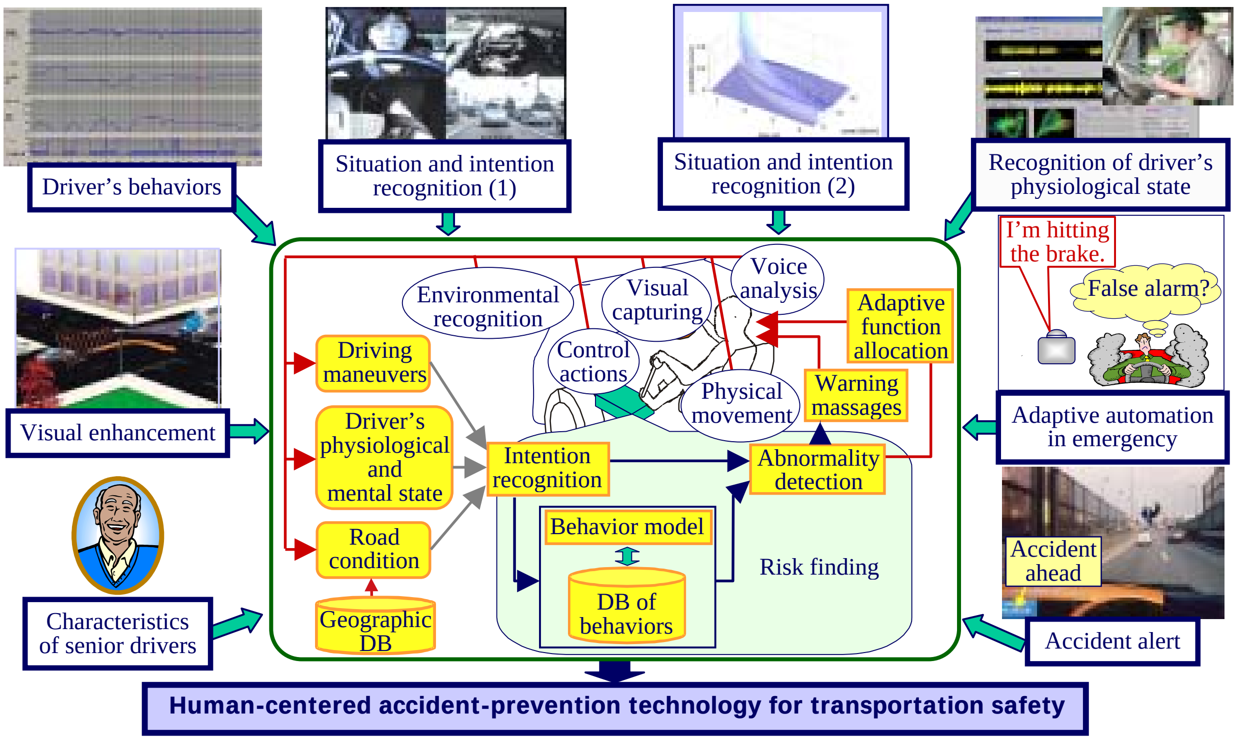


Figure 6: Situation and Intention Recognition for Risk Finding and Avoidance