International Journal of Vehicular Technology

Advances of Human Factors Research for Future Vehicular Technology

Guest Editors: Motoyuki Akamatsu, Paul Green, and Klaus Bengler

Advances of Human Factors Research for Future Vehicular Technology

International Journal of Vehicular Technology

Advances of Human Factors Research for Future Vehicular Technology

Guest Editors: Motoyuki Akamatsu, Paul Green,
and Klaus Bengler

Copyright © 2013 Hindawi Publishing Corporation. All rights reserved.

This is a special issue published in "International Journal of Vehicular Technology." All articles are open access articles distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Editorial Board

Sanghyun Ahn, Korea
Abdelaziz Bensrhair, France
Luis M. Bergasa, Spain
María Calderon, Spain
Tang-Hsien Chang, Taiwan
Kwang Cheng Chen, Taiwan
Jessie Y. C. Chen, USA
Sei-Wang Chen, Taiwan
Radu Danescu, Romania
Per Gårder, USA
Hilario Gómez-Moreno, Spain
Shinsuke Hara, Japan
Nicolas Hautière, France
Shin Kato, Japan
Alain Lambert, France
Xiaomin Ma, USA
P. T. Mathiopouloss, Greece
Rakesh Mishra, UK
Aboelmagd Noureldin, Canada
A. D. Panagopoulos, Greece
Nandana Rajatheva, Finland
C. S. Shankar Ram, India
Martin Reisslein, USA
Humphrey Rutagemwa, Canada
Fortunato Santucci, Italy
Sandro Scalise, Germany
Lingyang Song, China
Christian Wöhler, Germany
Yu-Dong Yao, USA
Kwan L. Yeung, Hong Kong

Contents

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2013, Article ID 749089, 2 pages
http://dx.doi.org/10.1155/2013/749089

Editorial

Advances of Human Factors Research for Future Vehicular Technology

Motoyuki Akamatsu,[1] Paul Green,[2] and Klaus Bengler[3]

[1] *Human Technology Research Institute, AIST, Japan*
[2] *University of Michigan Transportation Research Institute (UMTRI), USA*
[3] *Institute of Ergonomics, Technische Universität München, Germany*

Correspondence should be addressed to Motoyuki Akamatsu; akamatsu-m@aist.go.jp

Received 20 August 2013; Accepted 20 August 2013

Copyright © 2013 Motoyuki Akamatsu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Although automotive human factors research began after the World War II, vehicular technology has developed to adapt the vehicle to the human operator and the requirements of traffic since its inception, with the initial focus being on ease of operation of the steering wheel and brake pedal and methods to provide adequate road illumination at night. For many decades, human factors research mainly concerned making the primary driving tasks (controlling a vehicle and obeying signs and signals) easy to do, providing adequate space for the driver and passengers, mitigating crash injuries, and making secondary controls and displays inside the vehicle easy to use.

With the introduction of advanced driver assistance systems and driver information systems in 1980s, there have been a marked increase in the number of studies of driver mental workload as well as more general, quantitative studies of driver behavior, both on real roads and in driving simulators to help design and evaluate those systems. That line of human factors research will continue as vehicle automation and driver information increases. Another line of research concerns driver distraction, with a special concern being the use of mobile devices such as cellular phones. The implementation of driver assistance and information systems has resulted in a shift in automotive human factors research.

Methods for the automotive human factors research have been mostly transferred from psychology and cognitive science, physiology, statistics, and various engineering disciplines. But, as vehicular technology continues to evolve, new methods and theories are needed to address those issues and the human aspects of the vehicular technology so vehicles will be safe, easy to use, and useful.

The aim of the special issue is to collect research activities pertaining to human factors issue in the advanced driver assistance systems and driver information systems and to develop new methods for future automotive human factors issues.

The review article "*Automotive technology and human factors research: past, present, and future*" describes the history of automotive human factors research since the inception of motor vehicles until now. This article covers not only the research topics examined but also the industrial standards developed as a result of that research, the major organizations that research and conferences established at which it was presented, the major news stories affecting vehicle safety, and the general social context. It also identifies the issues to be addressed in the future.

We have two articles regarding the interface design of driver information systems. "*Development and evaluation of automotive speech interfaces: useful information from the human factors and the related literature*" by V. E. W. Lo, and P. A. Green provides a comprehensive literature review of the topic, summarizing much of what has been done in tables. Speech interfaces, though not commonly used now, have the potential of being less distracting to the driver than the visual-manual interfaces. The review article gives the background information such as key speech interfaces (both demonstrations and products), relevant technology standards and guidelines, and assessment methods for future development. "*Evaluation of a navigation radio using the think-aloud method*" by P. A. Green and J.-S. Park is a research article that describes how problems associated with a preproduction speech interface of a driver information system were identified using the think-aloud method commonly applied for usability testing. In addition to providing an extensive list of problems (common to many speech interfaces) as well as their frequency and severity, this article provides suggestions as to how the think-aloud method can be improved in this context so data can be collected and analyzed more quickly at a lower cost.

We have two research articles for assessing driver's capability of the driving task based on driving performance. "*Analysis of temporal relationships between eye-gaze and peripheral vehicle behavior for detecting driver distraction*" by T. Hirayama et al. gives a new measure to assess the driver distraction. They focused on the gaze behavior towards the peripheral vehicles and see how it changes depending on the level of distraction. "*A neurofuzzy approach to modeling longitudinal driving behavior and driving task complexity*" by R. G. Hoogendoorn, et al. focuses on the performance of the longitudinal driving task and how it changes depending on the driving task complexity, based on the theoretical framework of Task-Capability-Interface model that goes well beyond the classical Gazis-Herman-Rothery and Intelligent Driver models in the literature. They developed a mathematical model that can be used to assess the driver capability.

Overtrust and overreliance are key concerns when introducing advanced driver assistance systems but they are not clearly understood. "*Human's overtrust in and overreliance on advanced driver assistance systems: a theoretical framework*" by T. Inagaki and M. Itoh gives a two-aspect framework for the system (system diagnosis, action selection). System diagnosis is used to determine overtrust and has three characteristics (dimension of trust, target object, and chances for observation). The second aspect is used to determine overreliance and has three characteristics (type of action selected, benefits expected, and time allowance for intervention). Numerous examples of how these ideas can be applied to design real systems are given.

We have two research articles where the authors are developing methods for the human factors research. "*Predicting driver behavior using field experiment data and driving simulator experiment data: assessing impact of elimination of stop regulation at railway crossings*" by T. Sato et al. showed how the combined use of a naturalistic driving experiment and a driving simulator experiment can predict how drivers brake and where they stop if stop signs are removed from railway crossings. Of particular concern was the situation where vehicles ahead stop and the driver is trapped in the crossing and possibly killed by an oncoming train. "*The front seat passenger: how to transfer qualitative findings into design*" by S. Osswald et al. describes how qualitative methods can be used to identify ideas for new vehicle design, methods that contrast with traditional, quantitative methods used to test hypotheses using statistics.

"*A comparative analysis of subjective quality of the mobility between new portable electric transportation mode and walking*" by H. Ohta et al. is a research article concerning a Segway-type personal mobile vehicle. They investigated the subjective quality of mobility related to the use of this type of vehicle in different situations, to determine when it would be used and when people would prefer to walk or drive a car.

The variety of topics and methods described is indicative of the current state of the automotive human factors literature. For next generation motor vehicles and vehicle systems to be developed in a timely and cost-effective manner that customers will find useful, easy to use, and safe, efforts to (1) develop new methods, (2) collect and summarize the literature, (3) develop standards, guidelines, and recommended practices, and (4) evaluate prototype product designs must occur in parallel with human-centered design.

Motoyuki Akamatsu
Paul Green
Klaus Bengler

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2013, Article ID 285927, 8 pages
http://dx.doi.org/10.1155/2013/285927

Research Article

Analysis of Temporal Relationships between Eye Gaze and Peripheral Vehicle Behavior for Detecting Driver Distraction

Takatsugu Hirayama, Kenji Mase, and Kazuya Takeda

Graduate School of Information Science, Nagoya University, Nagoya 464-8601, Japan

Correspondence should be addressed to Takatsugu Hirayama; hirayama@is.nagoya-u.ac.jp

Received 24 September 2012; Revised 25 June 2013; Accepted 5 August 2013

Academic Editor: Klaus Bengler

Copyright © 2013 Takatsugu Hirayama et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A car driver's cognitive distraction is a main factor behind car accidents. One's state of mind is subconsciously exposed as a reaction reflecting it by external stimuli. A visual event that occurs in front of the driver when a peripheral vehicle overtakes the driver's vehicle is regarded as the external stimulus. We focus on temporal relationships between the driver's eye gaze and the peripheral vehicle behavior. The analysis result showed that the temporal relationships depend on the driver's state. In particular we confirmed that the timing of the gaze toward the stimulus under the distracted state induced by a music retrieval task using an automatic speech recognition system is later than that under a neutral state while only driving without the secondary cognitive task. This temporal feature can contribute to detecting the cognitive distraction automatically. A detector based on a Bayesian framework using this feature achieves better accuracy than one based on the percentage road center method.

1. Introduction

Driver distraction is a diversion of attention away from activities critical for safe driving toward a competing activity [1] and is a large risk factor that causes accidents [2]. Note that distraction differs from fatigue [3] which is defined as a state that disables one from continuing the activity [4]. Many researchers have developed driver distraction monitoring systems to maintain safety while driving by considering different types and levels of distraction [3]. The National Highway Traffic Safety Administration (NHTSA) classifies distractions into (1) cognitive distraction, (2) visual distraction, (3) auditory distraction, and (4) biomechanical distraction from the viewpoint of the driver's functionality [2]. Cognitive distraction can be considered as an internal state of the driver. It is difficult to sense this from outside. The other distractions are external factors that disturb the activity and can be observed more easily. We focus on cognitive distraction and seek novel findings to automatically detect it.

In the past few decades, a number of methods for detecting distraction have been proposed [3]. The methods fall into the following five categories based on the types of measures: (1) subjective report measures, (2) driver biological measures, (3) driving performance measures, (4) driver physical measures, and (5) hybrid measures. Among these measures, subjective report measures and driver biological measures are not suitable under real driving conditions. Driving performance measures as indicated by steering, braking behavior, and so forth are suitable for detecting visual distraction [5]. Even if a system can detect such overt behaviors that are more directly linked with risk, for maintaining safety, the timing may be too late to provide support to the driver after the detection.

Eye-gaze measure, which is one of the driver's physical measures, is a useful measurement of visual distraction especially for In-Vehicle Information System (IVIS) and Advanced Driver Assistance System (ADAS) assessment as specified in the existing standards ISO 15007-1 [6] and ISO/TS 15007-2 [7], and it has the potential for capturing symptoms of cognitive distraction [8]. An eye-gaze pattern could be used to discriminate driving while performing a secondary cognitive task from driving only [9]. Drivers under cognitive distraction had fewer saccades per unit time, which was consistent with less exploration of the driving environment [10]. Saccades may be a valuable index of mental workload [11]. Miyaji et al. reported that the standard deviations of eye movement and head movement could be suitable for detecting cognitive distraction that caused gaze concentration and slow saccades when drivers looked at the roadway [12]. Kircher et al. indicated the percentage of time that the driver spent observing the road ahead, which is called the percentage road center (PRC) of gaze direction, was more than 92% under cognitive distraction in a field study [13]. Johansson et al. have reviewed the existing gaze-based techniques and metrics for analyzing visual and cognitive distractions [8].

These approaches on driver's physics mainly measured only driver's gaze toward the road ahead or in-vehicle static objects without any regard for the peripheral traffic environment, which includes many scattering visual stimuli, or measured a rough correlation between spatial features of the gaze and the environment. They also need a long-term evaluation. To more flexibly support the driver, an improvement in the time resolution is required for the detection. We take account of short-term dynamics of cross media to detect cognitive distraction. In the field of human-computer interaction, some researchers have investigated a state of mind by analyzing the temporal relationships between eye movements and visual changes in the user interface [14–16]. The latent state is subconsciously exposed as a reaction reflecting it by external stimuli [17]. In controlled settings such as using a driving simulator, the detection response task (DRT) is an upcoming method of measuring visual and cognitive distractions [18–21], which asks the subject to respond via a device such as a button to visual, tactile, or acoustic stimuli. The response time relates to the distraction. However, it is difficult to give actual drivers on the road the task without disturbing the safety driving. In this work, our target is the temporal relationships between driver's gaze and peripheral vehicle behavior in a real driving situation. In particular it is the timing of the gaze toward the visual stimuli caused by the peripheral vehicle.

2. Analysis of Temporal Relationships between Driver Gaze and Peripheral Vehicle Behavior

2.1. Timing of Gaze Reaction to Overtaking Event. To analyze the temporal relationships, we focus on peripheral vehicle behaviors with a high level of visibility for the driver. When a peripheral vehicle (called the overtaking vehicle) overtakes the host vehicle driven by the driver, a visual change that occurs in front of the field of view can attract the driver's attention. We define this overtaking event as our target of analysis. The event has a base-point time $t_0(= 0)$, a beginning time $t_b(= t_0 - T_d/2)$, and an ending time $t_e(= t_0 + T_d/2)$. t_0 is the time when the front position y_o of the overtaking vehicle in the direction of forward movement becomes equal to the front position y_h of the host vehicle. T_d is the duration of the overtaking event, which is a configuration parameter of the analysis and set in Section 4.2. Figure 1 shows the overtaking event.

We define saccade timing t_c and gaze timing t_g. The former is the time when the driver turns gaze toward the overtaking vehicle, whereas the latter is the time while the driver fixates the overtaking vehicle. The temporal relationships characterizing the gaze reaction to the overtaking vehicle are the time differences between the saccade timing t_c or the gaze timing t_g and the base-point time t_0 of the event. Figure 2 shows the timing structure.

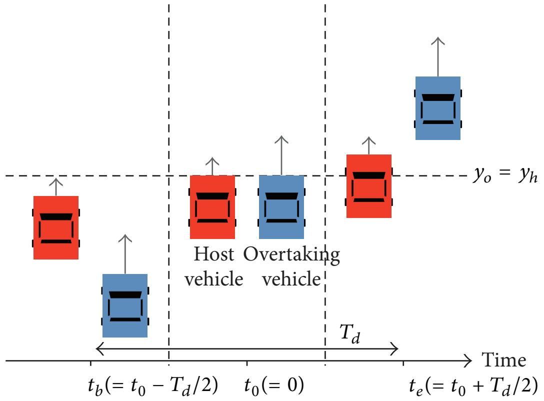


Figure 1: Overtaking event.

2.2. Hypothesis. Some researchers have investigated the correlations between gaze directions and traffic object positions, for example, road curvature, oncoming traffic, road signs, and pedestrians [22–24], and have shown that the correlations are high under the neutral state while driving only. Most of the researchers, however, dealt mainly with spatial correlation. On the temporal relationship between visual attention and external stimuli, Posner revealed that covert attention to them decreases the reaction time, and, conversely, distraction increases it [25]. We therefore propose the following hypothesis: *the timing of when a driver gazes toward the overtaking event under a state of cognitive distraction is later than that under a neutral state.*

3. Real-World Driving Database

We analyze a part of a database collected using the "NUDrive Vehicle" in Nagoya, Japan [26].

3.1. Data-Collection Vehicle. The "NUDrive Vehicle" was designed to synchronously record multimedia signals of driver performance (gas pedal, brake pedal, steering angle, velocity, acceleration, and position of the car), intervehicular distance, biological signals, videos, and audio signals. Various external sensors were mounted on a Toyota Hybrid Estima with a 2360 cc engine and automatic transmission and steering wheel on the right side. Figure 3 shows the data-collection vehicle. All the sensors used for recording were commercially available.

3.2. Participants. A total of 30 participants (10 males and 20 females) took part in the experiment. They were, on average, 39.0 years old (range of 29–52 years) and had held a driver's license for a mean period of 18.6 years (range of 8–32 years). They received 5000 Japanese yen as compensation for their participation.

3.3. Procedure. The participants first drove for a few minutes to get used to the vehicle and the sensors. Signals recorded

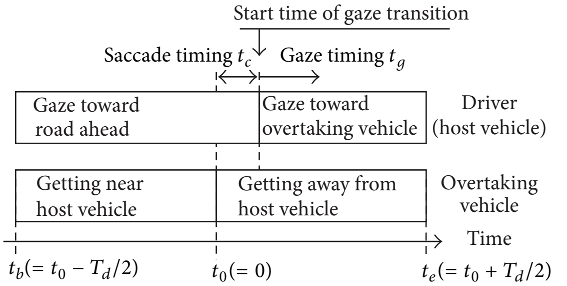


Figure 2: Timing of the driver's gaze toward the overtaking vehicle.

Figure 3: Data-collection vehicle.

during the initial period were not used in this work. The experimenter monitored the experiment from the rear seat and indicated the route to the driver. During a particular period of driving, the participant performed a secondary hands-free task of retrieving and playing songs from a list of 635 titles from 248 artists using an automatic speech recognition system [27]. The secondary task artificially induced the cognitive distraction state. The experimenter instructed the participant to retrieve as many songs as possible; accordingly, within around 30 s of successfully retrieving each song, the participant had to retrieve another song. All experiments were performed on two-or three-lane highways. The experimental route was the same for all participants.

3.4. Measures. In this work, we analyzed the intervehicular distance measured by laser scanners and the driver's gaze direction extracted from the recorded video. The following are the details of the analyzed data and omit the account of other data.

3.4.1. Intervehicular Distance Scanning. Two laser scanners (front: RIEGL LMS-140i-80; rear: RIEGL LMS-Q120i), mounted on the front and back of the host vehicle shown in Figure 3, provided geometric information about the peripheral environment of the vehicle. The laser scanners covered 80-degree arcs at both front and back of the vehicle, to an effective range of about 100 m to the front and 55 m to the rear, but had blind areas at the left and right sides of the vehicle. The data were acquired at a sample frequency of 10 Hz. For tracking peripheral vehicles in the blind areas, we applied a Kalman filter to the data [28]. The dynamics of their position and velocity relative to the host vehicle could be estimated even if they were outside the laser range. The position was on a horizontal plane whose coordinate system was comprised of a moving directional axis y and its orthogonal axis x with origin (x_0, y_0) at the center of the frontal laser scanner. We did not take velocity into account in this work. The practical area to analyze is limited to a rectangular area with a length of 80 m, $-40 \le y \le 40$, and a width of 9.9 m, $-4.95 \le x \le 4.95$.

3.4.2. Video Recording. The driver's face was captured by a camera (Sony 1/2 inch CCD video camera DXC-200A) mounted on the dashboard. The data were acquired at a resolution of 692 pixels in width and 480 pixels in height at a sample frequency of 29.4 fps.

3.4.3. Driver Gaze Tracking. The driver's gaze direction was manually labeled using ELAN, (http://www.lat-mpi.eu/tools/elan) which is a tool for the creation of complex annotations on video and audio resources, by an annotator. We prepared five gaze labels according to ISO 15007-1 [6] and three additional labels to detect gaze toward overtaking vehicles and toward upper traffic signs that could be extracted from the low-resolution video as follows:

(g_0) *right side* (gaze toward right mirror and right window by head turning);

(g_1) *right front* (gaze rightward from *front*, including gaze toward overtaking vehicles in the right lane);

(g_2) *rear* (gaze toward rear-view mirror);

(g_3) *front* (gaze to road scene ahead, the reference direction);

(g_4) *left side* (gaze toward left mirror and left window by head turning);

(g_5) *left front* (gaze leftward from *front*, including gaze toward overtaking vehicles in the left lane);

(g_6) *up* (gaze upward from front, including gaze toward upper traffic signs);

(g_7) *down* (gaze downward from *front*, including gaze toward instrument panel).

The annotator detected the beginning of the saccade of each gaze behavior as the beginning of interval with the gaze label. We consider that the gaze direction can be labeled more stably and accurately by using commercially supplied eye-tracking systems (e.g., FaceLab). Figure 4 shows a sample set of face images that were given each label.

4. Timing Analysis of Driver Gaze under Cognitive Distraction toward Peripheral Vehicle Behavior

4.1. Extracting Overtaking Events. We extracted 274 overtaking events from the intervehicular distance data of the neutral task (task condition $\mathscr{C}_N$) and 81 events from that of the music-retrieval task (task condition $\mathscr{C}_M$).

(a) Left-front (b) Front (c) Up (d) Right-front

(e) Left-side (f) Rear (g) Down (h) Right-side

FIGURE 4: Sample set of gaze labels and face images.

4.2. Clustering of Peripheral Environment Data. We do not analyze traffic simulation data, we but focus on real traffic data. There are some peripheral vehicles in addition to the overtaking vehicle that act as triggers of random events in a real traffic scene. The peripheral vehicles are the visual stimuli. From the standpoint of the visual target search task, the overtaking vehicle and the other peripheral vehicles are regarded as target and visual distractors, respectively. The performance of the target search task depends on the traits of the visual distractors [29]. In this work, a trait of the peripheral environment needs to be defined. We classify the traffic scenes spatiotemporally according to time while the peripheral vehicles exist in the subperipheral area within the interval of the overtaking event, T_d. The peripheral area is divided into six subareas corresponding to the six gaze labels g_j $(j = 0, \ldots, 5)$. Figure 5 shows the six subperipheral areas. Each of the six areas is assigned a Boolean variable a_j. A vector, $\mathbf{e}$, comprises the six variables: $\mathbf{e} = (a_5, a_4, a_3, a_2, a_1, a_0)$, which represents the trait of the peripheral environment. In this work, if the cumulative duration while the peripheral vehicles exist in subperipheral area a_j goes over 50% of the targeted time interval, $a_j = 1$; otherwise, $a_j = 0$. We focus on the dynamics of the driver's gaze on the base point of the overtaking event, t_0, as a reference point of analysis. Therefore, the interval of the overtaking event, T_d, is divided into the first half $\mathbf{e}_1$ from t_b to t_0 and the second half $\mathbf{e}_2$ from t_0 to t_e. Practically, we set T_d to 10 s ($t_0 = 0, t_b = -5, t_e = 5$(s)) as a sufficient time duration to analyze in consideration of the analysis area with the length of 80 m and the maximum velocity difference (=40.9 km/h) between the host vehicle and the overtaking vehicles in the area.

From all data including the overtaking events: 274 events for $\mathscr{C}_N$ and 81 events for $\mathscr{C}_M$, we could not extract uniformly distributed vectors $\mathbf{e}_1$ and $\mathbf{e}_2$. In addition, we eliminated some traffic scenes along curves and including lane changes because they would induce a specific gaze behavior. Adequate samples were extracted for two types of state transitions of peripheral environment as follows: environmental state transition $\mathscr{E}_A$ from $\mathbf{e}_1 = (0, 0, 0, 0, 0, 1)$ to $\mathbf{e}_2 = (0, *, 0, *, 1, *)$, 75 events for task $\mathscr{C}_N$ and 23 events for task $\mathscr{C}_M$, and environmental state transition $\mathscr{E}_B$ from $\mathbf{e}_1 = (0, 0, 0, 0, 1, 1)$ to $\mathbf{e}_2 = (0, *, 0, *, 1, *)$, 43 events for task $\mathscr{C}_N$ and 13 events for task $\mathscr{C}_M$. 43.0% for $\mathscr{C}_N$ and 44.4% for $\mathscr{C}_M$ of all overtaking events were fallen into either $\mathscr{E}_A$ or $\mathscr{E}_B$. "$*$" denotes any binary value. The former represents when the overtaking vehicle runs in the right lane and the other peripheral vehicles do not exist for more than 2.5 s (=50% of $T_d/2$) before it overtakes the host vehicle, whereas the latter represents when the overtaking vehicle also runs in the right lane and a peripheral vehicle exists in the *right-front* area for more than 2.5 s before it overtakes the host vehicle. We analyze the relationships between the extracted events for two environmental conditions $\mathscr{E}_A$ and $\mathscr{E}_B$ and the gaze data below.

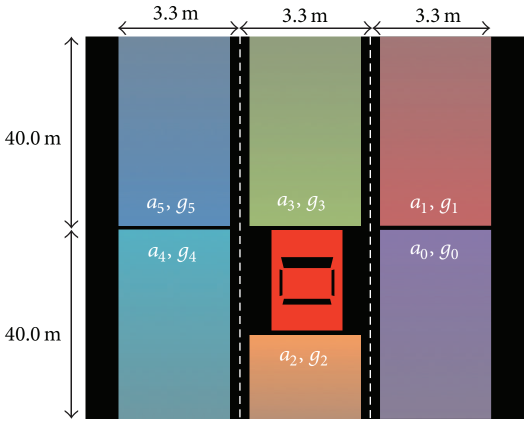
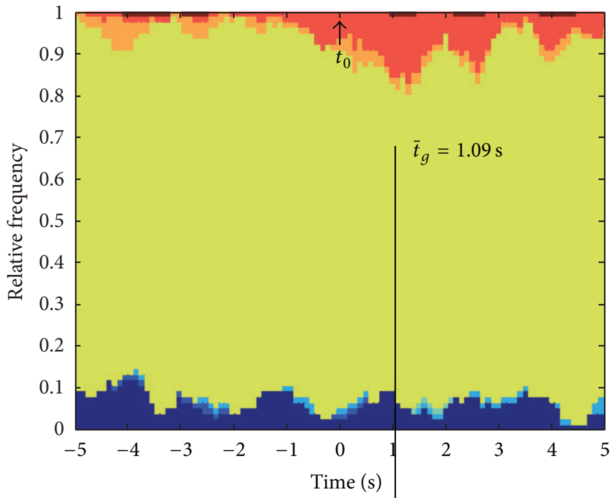
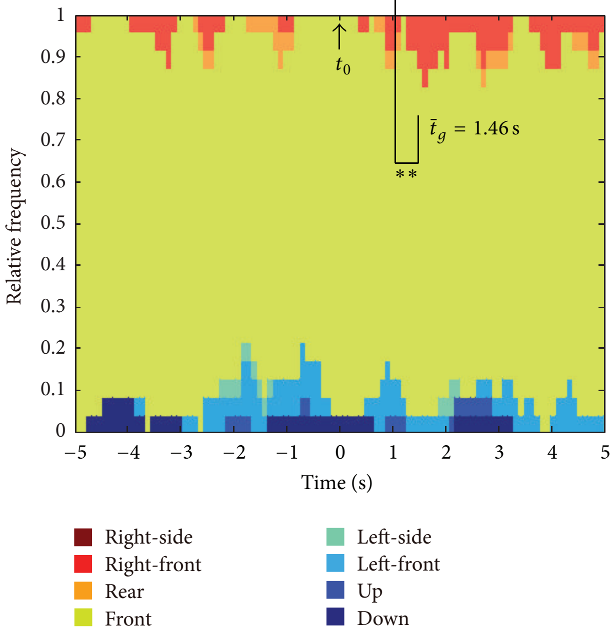


FIGURE 5: Definition of six subperipheral areas.

4.3. Testing of Hypothesis Based on Temporal Gaze Distribution. The relative frequency distributions of the gaze directions (temporal gaze distributions) that were measured in the peripheral environment $\mathscr{E}_A$ varied in terms of time, as shown in Figure 6. Since the relative frequency of the *right-front* gazes increased quickly after the base point of the overtaking event, t_0, regardless of the task, the participants frequently gazed toward the overtaking vehicle in the right lane when the number of visual distractors was small. We verified



(a) Neutral state, $\mathscr{C}_N$



(b) Distraction state, $\mathscr{C}_M$

FIGURE 6: Time variations of relative frequency distributions of gaze directions that were measured in the peripheral environment $\mathscr{E}_A$.

the hypothesis (see Section 2.2) of the temporal difference between the relative frequency distributions for the music-retrieval task $\mathscr{C}_M$ and the neutral task $\mathscr{C}_N$ by performing a statistical test. Let us limit the interval for testing to the time from $t_s = t_0 (= 0)$ to $t_f = 2$(s) following the work of Merat et al. [19]. The average saccade timings $\bar{t}_c$ for the two conditions $\mathscr{C}_N$ and $\mathscr{C}_M$ were 0.82 s and 1.07 s (SD = 0.52 s and 0.42 s), respectively. The saccade timing for the $\mathscr{C}_N$ condition was shorter than that for the $\mathscr{C}_M$ condition, but we did not obtain adequate samples for performing any statistical test to the saccade timings. The average gaze timings $\bar{t}_g$ for the two conditions $\mathscr{C}_N$ and $\mathscr{C}_M$ were 1.09 s and 1.46 s (SD = 0.51 s and 0.43 s), respectively. The two relative frequency distributions within the limited interval had normality and homoscedasticity. The t-test revealed that the gaze timing for the $\mathscr{C}_N$ condition was significantly shorter than that for the $\mathscr{C}_M$ condition, $t(183) = -3.76$, $P < .001$.

In contrast, Figure 7 shows that the participants rarely looked at the overtaking vehicle in the peripheral environment $\mathscr{E}_B$ while performing the music-retrieval task $\mathscr{C}_M$, whereas they followed the same behavior as $\mathscr{E}_A$ for the neutral task $\mathscr{C}_N$. In the same manner as $\mathscr{E}_A$, the average gaze timings $\bar{t}_c$ for the two conditions $\mathscr{C}_N$ and $\mathscr{C}_M$ were 0.88 s and 1.25 s (SD = 0.67 s and 0.60 s), respectively. The saccade timing for the $\mathscr{C}_N$ condition was shorter than that for the $\mathscr{C}_M$ condition, but we did not obtain adequate samples for performing any statistical test to the saccade timings. The average gaze timings $\bar{t}_g$ for the two conditions $\mathscr{C}_N$ and $\mathscr{C}_M$ were 0.93 s and 1.34 s (SD = 0.51 s and 0.45 s), respectively. The t-test revealed that the gaze timing for the $\mathscr{C}_N$ condition was significantly shorter than that for the $\mathscr{C}_M$ condition, $t(83) = -3.1$, $P < .005$. These results support our hypothesis but are limited to evaluation for two conditions, $\mathscr{E}_A$ and $\mathscr{E}_B$, of the peripheral environment.

5. Detection of Driver Distraction

5.1. Discrimination between Distraction and Neutral State in a Bayesian Framework. To identify the class label (*distraction*, i.e., $\mathscr{C}_M$, or *neutral*, i.e., $\mathscr{C}_N$) of the driver state using the temporal gaze distribution, we use a naive Bayesian framework as follows:

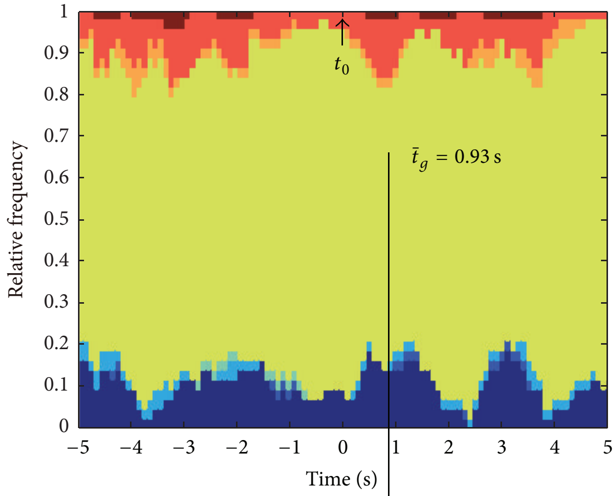
$$P(C_t \mid G_t, \mathbf{E}_t) = \frac{P(C_t, \mathbf{E}_t)\, P(G_t \mid C_t, \mathbf{E}_t)}{P(G_t, \mathbf{E}_t)} = \frac{P(C_t \mid \mathbf{E}_t)\, P(\mathbf{E}_t)\, P(G_t \mid C_t, \mathbf{E}_t)}{P(G_t, \mathbf{E}_t)}, \quad (1)$$

where t represents the time from when the host vehicle was overtaken by the other vehicle, C_t represents whether a gaze belongs to the distraction class at time t, that is, the binary class label, G_t represents whether the direction of the gaze is *right-front*, that is, gaze label, and $\mathbf{E}_t$ represents the condition of the peripheral environment. One of the important characteristics of the Bayesian framework is the capability to infer the state of an unobserved variable, given the state of the observed variables. In our case, we want to infer the driver's internal state, that is, cognitive distraction, given the gaze data and the peripheral environment.

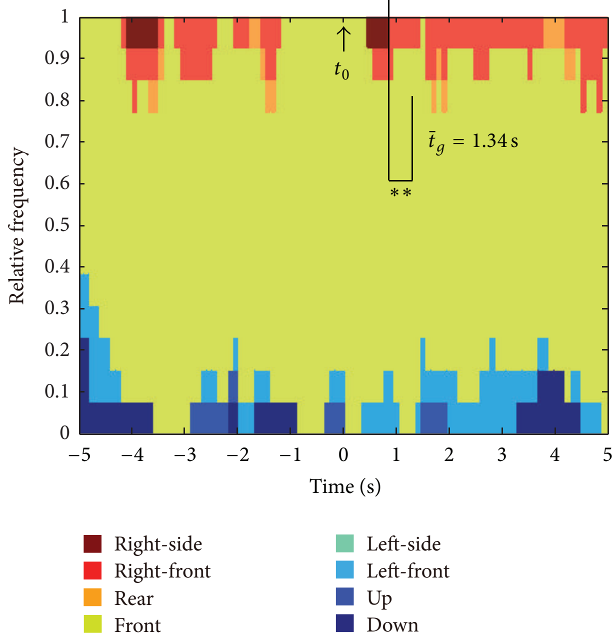
To make a decision as to the class of discrimination is assigned to the gaze data, the equation can be iterated over a time interval related to an overtaking event. We can then accumulate the computed posteriors and choose the class of driver state with greater score based on Maximum a Posteriori (MAP) as follows:

$$Q(C \mid G, \mathbf{E}) = \sum_{t=t_s}^{t_f} P(C_t \mid G_t, \mathbf{E}_t), \quad (2)$$

$$S = \arg\max_{C} Q(C \mid G, \mathbf{E}). \quad (3)$$



(a) Neutral state, $\mathscr{C}_N$



(b) Distraction state, $\mathscr{C}_M$

FIGURE 7: Time variations of relative frequency distributions of gaze directions that were measured in the peripheral environment $\mathscr{E}_B$.

In this work, the priors $P(C_t \mid \mathbf{E}_t)$ and $P(\mathbf{E}_t)$ are assumed as uniform, that is, noninformative prior distributions, and C_t and $\mathbf{E}_t$ are regarded as binary variables not depending on time t: $P(G_t \mid C, \mathbf{E}) \simeq P(G_t \mid C_t, \mathbf{E}_t)$, $\mathbf{E} = (\mathbf{e}_1, \mathbf{e}_2)$. Therefore, we can apply the temporal gaze distribution extracted in Section 4.3 to $P(G_t \mid C, \mathbf{E})$.

5.2. Results. We matched the class label computed from equation (3) with the true label to discriminate between distraction state $\mathscr{C}_M$ and neutral state $\mathscr{C}_N$. The experimental data were the same as the data for analysis of gaze timing (see Section 4): 75 events for $\mathscr{C}_N$ and 23 for $\mathscr{C}_M$ in $\mathscr{E}_A$, 43 events for $\mathscr{C}_N$, and 13 for $\mathscr{C}_M$ in $\mathscr{E}_B$. We applied leave-one-out cross validation to obtain the discrimination accuracy.

Table 1 shows the accuracies of the two-class discrimination. As this is for testing our hypothesis (see Section 4.3), we limited the time interval for discrimination to from $t_s = t_0 (= 0)$ to $t_f = 2$(s). Here, we employed a baseline method based on the percentage road center (PRC) [13], that is, proportion of total gaze duration toward road scene ahead to total time of sample events, which detected $\mathscr{C}_M$ and $\mathscr{C}_N$ by thresholding PRC at π_r; that is, if PRC was larger than π_r, the method determined the data as distraction state $\mathscr{C}_M$; otherwise, neutral state $\mathscr{C}_N$ was assigned. The thresholds π_r for $\mathscr{E}_A$ and $\mathscr{E}_B$ were set to 79.0%, by searching for an equal rate between detection of $\mathscr{C}_M$ and $\mathscr{C}_N$. We can confirm that the proposed method performed more accurately than the baseline one.

6. Discussion

We obtained better test results supporting our hypothesis and confirmed that the proposed discriminator performed more accurately than the baseline one. This approach takes advantage of the shorter time needed to detect cognitive distraction because it focuses on only the important scene for the detection but needs to trigger the gaze reaction to the overtaking event. The Bayesian rule-based method excels in application. It can be naturally integrated into the state-of-the-art method based on Bayesian networks using hybrid measures [30].

Figures 6 and 7 also suggest another difference between two tasks $\mathscr{C}_M$ and $\mathscr{C}_N$. Note the frequency of gazing to a downward direction, that is, *down*. We consider that the participants frequently looked at the speedometer or navigation system while performing the neutral task. This behavior agrees with the prior findings of PRC [13]. In the temporal section without the overtaking event, PRC needs to be addressed to detect cognitive distraction.

Here, let us compare the average timing of the gaze to *right front* in condition $\mathscr{E}_A$ with that in condition $\mathscr{E}_B$ (see their values in Section 4.3). We can confirm that the latter was slightly shorter than the former. In condition $\mathscr{E}_B$, there was a vehicle that preceded the overtaking vehicle. The participants might still focus their attention on the preceding vehicle or *right-front* area and then effectively react to the next overtaking event. The event did not cause inhibition of return [31], which retards their reaction.

These results were verified under only two limited environmental conditions, that is, $\mathscr{E}_A$ and $\mathscr{E}_B$, of peripheral vehicles because we could not analyze enough experimental data. We need to increase the number of clusters of peripheral environment data for wide-ranging analysis and to model the dynamics of peripheral vehicles based on, for example, time to collision (TTC), velocity, acceleration, and interaction among vehicles for deeply analyzing the temporal relationships and increasing the discrimination accuracy. We also have to analyze the other secondary tasks and differences among individuals.

7. Conclusions

The dynamics of the external environment can elicit reactions reflecting the human internal state, that is, make the latent

Table 1: Two-class discriminant accuracies.

	Peripheral environment $\mathscr{E}_A$		Peripheral environment $\mathscr{E}_B$	
	Distraction ($\mathscr{C}_M$)	Neutral ($\mathscr{C}_N$)	Distraction ($\mathscr{C}_M$)	Neutral ($\mathscr{C}_N$)
Proposed method	66.7%	64.7%	75.0%	61.5%
Baseline method (PRC)	50.0%	52.9%	50.0%	53.8%

state explicit. We showed that the temporal factor, that is, timing, of a reaction is important for understanding the state by focusing on cognitive distraction in a car-driving situation. The concrete contribution of this paper is twofold. First, we obtained test results supporting our hypothesis that the timing of when a driver gazes toward the overtaking event under cognitive distraction is later than that under the neutral state. Second, we confirmed that a Bayesian-based detection of distraction using the temporal gaze distribution performed more accurately than the PRC-based one. The findings of this work should be generalized through additional analysis in future work. We have built a large database of 500 drivers [26]. The generalized findings will suggest a risk of voice interactive navigation (hands-free navigation) using automatic speech recognition and a novel testing scenario without driver's extra workload for driver information system in real driving situation to the Alliance of Automobile Manufacturers (AAM) guidelines [32]. To put our approach into practical use, we will refer to peripheral vehicle-tracking systems based on computer vision techniques [24, 33] and replace the laser scanner with a camera-based system.

Acknowledgments

This work is partially supported by JST CREST (Core Research for Evolutional Science and Technology of Japan Science and Technology Agency) Kashin-free project and a Grant in Aid for Scientific Research from MEXT (Ministry of Education, Culture, Sports, Science, and Technology) of Japan under Contract no. 23700168. The authors thank Professor Chiyomi Miyajima and Mr. Masataka Mori for collecting experimental data.

References

[1] M. A. Regan, J. D. Lee, and K. L. Young, "Defining driver distraction," in *Driver Distraction: Theory, Effects, and Mitigation*, chapter 4, pp. 42–54, CRC, 2008.

[2] T. A. Ranney, W. R. Garrott, and M. J. Goodman, *NHTSA Driver Distraction Research: Past, Present, and Future*, National Highway Traffic Safety Administration, 2001.

[3] Y. Dong, Z. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: a review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 596–614, 2011.

[4] H. Croo, M. Bandmann, G. Mackay, K. Rumar, and P. Vollenhoven, *The Role of Driver Fatigue in Commercial Road Transport Crashes*, 2001.

[5] Y. Liang and J. D. Lee, "Combining cognitive and visual distraction: less than the sum of its parts," *Accident Analysis and Prevention*, vol. 42, no. 3, pp. 881–890, 2010.

[6] ISO, 15007-1:2002, "Road vehicles—measurement of driver visual behavior with respect to transport information and control systems—part 1: definitions and parameters," 2002.

[7] ISO/TS, 15007-1:2001, "Road vehicles—measurement of driver visual behavior with respect to transport information and control systems—part 2: equipment and procedures," 2001.

[8] E. Johansson, J. C. Engström, C. Cherri et al., *Review of existing techniques and metrics for IVIS and ADAS assessment*, Adaptive Integrated Driver-Vehicle Interface, 2004.

[9] L. Angell, J. Auflick, A. Austria et al., *Driver Workload Metrics Project—Task 2 Final Report*, U.S. Department of Transportation, National Highway Traffic Safety Administration, 2006.

[10] J. L. Harbluk and Y. I. Noy, *The Impact of Cognitive Distraction on Driver Visual Behavior and Vehicle Control*, Ergonomics Division, Road Safety Directorate and Motor Vehicle Regulation Directorate, Ontario, Canada, 2002.

[11] J. G. May, R. S. Kennedy, M. C. Williams, W. P. Dunlap, and J. R. Brannan, "Eye movement indices of mental workload," *Acta Psychologica*, vol. 75, no. 1, pp. 75–89, 1990.

[12] M. Miyaji, H. Kawanaka, and K. Oguri, "Driver's cognitive distraction detection using physiological features by the AdaBoost," in *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems (ITSC '09)*, pp. 90–95, October 2009.

[13] K. Kircher, C. Ahlstrom, and A. Kircher, "Comparison of two eye-gaze based real-time driver distraction detection algorithms in a small-scale field operational test," in *Proceedings of the 5th International Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, pp. 16–23, 2009.

[14] R. Ishii and Y. I. Nakano, "Estimating user's conversational engagement based on gaze behaviors," in *Intelligent Virtual Agents, Lecture Notes in Computer Science*, vol. 5208, pp. 200–207, 2008.

[15] T. Hirayama, J. Dodane, H. Kawashima, and T. Matsuyama, "Estimates of user interest using timing structures between proactive content-display updates and eye movements," *IEICE Transactions on Information and Systems*, vol. E93-D, no. 6, pp. 1470–1478, 2010.

[16] R. Yonetani, H. Kawashima, T. Hirayama, and T. Matsuyama, "Mental focus analysis using the spatio-temporal correlation between visual saliency and eye movements," *Journal of Information Processing*, vol. 20, no. 1, pp. 267–276, 2012.

[17] T. Hirayama, Y. Sumi, T. Kawahara, and T. Matsuyama, "Info-concierge: proactive multi-modal interaction based on mind probing," in *The Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA/ASC '11)*, 2011.

[18] N. Merat and A. H. Jamson, "Multisensory signal detection: a tool for assessing driver workload during IVIS management," in *Proceedings of the 4th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2007.

[19] N. Merat, E. Johansson, J. A. Engström, E. Chin, F. Nathan, and T. W. Victor, *Specification of a secondary task to be used*

in safety assessment of IVIS, Adaptive Integrated Driver-Vehicle Interface, 2007.

[20] G. L. Rupp, *Performance Metrics for Assessing Driver Distraction: The Quest for Improved Road Safety*, SAE International, 2010.

[21] L. Hsieh, R. Young, and S. Seaman, "Development of the enhanced peripheral detection task: a surrogate test for driver distraction," *SAE International Journal of Passenger Cars*, vol. 5, no. 1, pp. 317–325, 2012.

[22] M. F. Land and D. N. Lee, "Where we look when we steer," *Nature*, vol. 369, no. 6483, pp. 742–744, 1994.

[23] N. Apostoloff and A. Zelinsky, "Vision in and out of vehicles: integrated driver and road scene monitoring," *International Journal of Robotics Research*, vol. 23, no. 4-5, pp. 513–538, 2004.

[24] L. Fletcher and A. Zelinsky, "Driver inattention detection based on eye gaze-road event correlation," *International Journal of Robotics Research*, vol. 28, no. 6, pp. 774–801, 2009.

[25] M. I. Posner, "Orienting of attention," *The Quarterly journal of experimental psychology*, vol. 32, no. 1, pp. 3–25, 1980.

[26] K. Takeda, J. H. L. Hansen, P. Boyraz, L. Malta, C. Miyajima, and H. Abut, "International large-scale vehicle corpora for research on driver behavior on the road," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 1–15, 2011.

[27] S. Hara, C. Miyajima, K. Itou, and K. Takeda, "An online customizable music retrieval system with a spoken dialogue interface," *Journal of Acoustical Society of America*, vol. 120, no. 5, pp. 3378–3379, 2006.

[28] Y. Li, C. Miyajima, N. Kitaoka, and K. Takeda, "Driving scene retrieval using integrated vehicle motion feature matching," in *Proceedings of the 5th Biennial Workshop on DSP for In-Vehicle Systems*, pp. 1–8, 2011.

[29] J. M. Wolfe and T. S. Horowitz, "What attributes guide the deployment of visual attention and how do they do it?" *Nature Reviews Neuroscience*, vol. 5, no. 6, pp. 495–501, 2004.

[30] L. Malta, C. Miyajima, N. Kitaoka, and K. Takeda, "Analysis of real-world driver's frustration," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 109–118, 2011.

[31] M. I. Posner and Y. Cohen, "Components of visual orienting," in *Attention and Performance*, X. H. . Bouma and D. Bowhuis, Eds., pp. 531–556, Lawrence Erlbaum Associates, Hillsdale, NJ, USA, 1984.

[32] Alliance of Automobile Manufactures, *Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems*, 2006.

[33] A. Doshi and M. Trivedi, "Investigating the relationships between gaze patterns, dynamic vehicle surround analysis, and driver intentions," in *Proceedings of the IEEE Intelligent Vehicles Symposium*, pp. 887–892, June 2009.

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2013, Article ID 526180, 27 pages
http://dx.doi.org/10.1155/2013/526180

Review Article

Automotive Technology and Human Factors Research: Past, Present, and Future

Motoyuki Akamatsu,[1] Paul Green,[2] and Klaus Bengler[3]

[1] *Human Technology Research Institute, AIST, Japan*
[2] *University of Michigan Transportation Research Institute (UMTRI), USA*
[3] *Institute of Ergonomics, Technische Universität München, Germany*

Correspondence should be addressed to Motoyuki Akamatsu; akamatsu-m@aist.go.jp

Received 14 February 2013; Accepted 7 May 2013

Academic Editor: Tang-Hsien Chang

Copyright © 2013 Motoyuki Akamatsu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper reviews the history of automotive technology development and human factors research, largely by decade, since the inception of the automobile. The human factors aspects were classified into primary driving task aspects (controls, displays, and visibility), driver workspace (seating and packaging, vibration, comfort, and climate), driver's condition (fatigue and impairment), crash injury, advanced driver-assistance systems, external communication access, and driving behavior. For each era, the paper describes the SAE and ISO standards developed, the major organizations and conferences established, the major news stories affecting vehicle safety, and the general social context. The paper ends with a discussion of what can be learned from this historical review and the major issues to be addressed. A major contribution of this paper is more than 180 references that represent the foundation of automotive human factors, which should be considered core knowledge and should be familiar to those in the profession.

1. Introduction

In many fields of technology, examinations of the past can provide insights into the future. This paper examines (1) the driver- and passenger-related technology that was developed as a function of time and (2) the research necessary for those developments, as they affected both vehicle design and evaluation. This paper also examines how those developments were influenced by (1) advances in basic technology, (2) requirements from government agencies and international standards, and (3) even the news media. All of this is done roughly chronologically, with developments grouped into three time periods—before World War II, after World War II until 1989, and since 1990.

In the history of research, a research topic becomes popular at some time because of a societal need, researcher interest, technology trends, the introduction of a new method, or a new theory. As a consequence, the number of researchers in the field grows, as does the number of publications, which in turn leads to products, services, and new ideas. These factors have certainly affected the growth of the human factors profession.

The history of automotive technology and human factors research can be viewed similarly. Its history can be divided into three periods. They are (1) the decades before World War II (Section 2), (2) World War II until 1989 (Section 3), and (3) 1990 and beyond (Section 4). This last period is continuing, so it is a bit more difficult to be retrospective in grouping decades. Therefore, Section 4 is divided by research topics, not by decades. For each topic, research activities are described chronologically to help readers to understand how the research has progressed for these 20 years to reach the current status.

2. A Short History of Human Factors Aspects of Automotive Technology before World War II

2.1. Early Stage of Automobiles (1886–1919). Over the course of the first half-century after the invention of the automobile

(a) (b)

Figure 1: Tiller (Oldsmobile 1902 (a)) and bar handle (Peugeot Type 24 1899 (b)) (the author's (MA) photo collection).

by Karl Benz in 1886, various changes were made to self-powered vehicles so they were better suited to human abilities, changes based on experience with animal-drawn vehicles. Interestingly, the seatbelt had been introduced for steam-powered horseless carriages in the 1800s, but its purpose was to keep passengers on their seat, not to keep them safe in the event of a collision [1]. The steering mechanism in very early automobiles was a tiller, a lever arm that connected to the pivot point of the front wheels, a design derived from small boats. Tillers were easy to use for very slow speeds and lightweight vehicles (such as those with three wheels). However, steering a jolting tiller with sheer muscle power was difficult for heavy four-wheel vehicles moving at high speed. A bar handle with grips at both ends to be held with both hands was introduced that could be held more firmly than the tiller. A round steering wheel, able to be turned by muscle power and easier to hold in the hands, was first introduced around 1895 (Figure 1).

The brake system for very early self-powered vehicles consisted of a wooden block pressed against one of the wheels using a hand-operated lever, a technology adapted from horse carriages. A foot pedal to operate the band brake first appeared in Benz Velo in 1894 (Figure 2). The foot-operated pedal could exert greater force than a hand brake and allowed a driver to use both hands to hold a steering wheel. This could be why the steering wheel and the foot pedal appeared in the same period.

Early automobiles were not equipped with any gauges. Oil-pump gauges were the first instruments installed inside vehicles, allowing drivers to confirm the oil flow and to inject additional oil when necessary. Water-pressure gauges were also introduced around 1900. Durability was the biggest issue in the early stage of automobiles. Therefore, general monitoring of the condition of unreliable vehicles by the driver was critical and consumed considerable attention.

The speedometer was introduced after 1900. It was mounted outside of the bulkhead separating the engine and cab, where its cable easily fits. The speedometer was introduced to highlight the vehicle's high-speed capability. In the USA, the state of Connecticut imposed a speed limit of 8 mph within the city and 12 mph outside of the city in 1901, thus encouraging speedometer installation [2].

The manufacturer Panhard et Levassor first placed a radiator in the front end of the vehicle for effective cooling. A thermometer was installed on top of the radiator in the early 1910s, allowing the driver to read the temperature from the driver's seat. Making sure the instrument was visible to the driver and was easy to install were important design considerations. In many cases, the hood ornament on contemporary vehicles is a remnant of these instruments.

After around 1910, instruments such as tachometers and clocks were installed inside automobiles. These were directly fixed on the surface of the bulkhead, and visibility to the driver was poor (Figure 3(a)). In the late 1910s, instrument panels (or dashboards) were installed separately from the bulkheads (Figure 3(b)). The instrument panel configurations were inconsistent. Some manufacturers concentrated the gauges in the central area of the panel and others distributed them across the panel.

An indication of the importance of the industry was the growth of organizations to support it. In 1901, the later German Verband der Automobilindustrie (VDA) association of automotive industry was founded as Verein Deutscher Motorfahrzeug-Industrieller (VDMI). VDMI was established to promote road transport, defend against "burdensome measures by the authorities" (taxation, liability obligations), support customs protection, and monitor motor shows. In 1923, the VDMI was renamed the Reichsverband der Automobilindustrie (RDA). The present name Verband der Automobilindustrie (VDA) was given to this umbrella organization of the German automotive industry in 1946 (http://www.vda.de/en/verband/historie.html).

To exchange engineering ideas to facilitate the growth of the automotive industry, the Society of Automotive Engineers (SAE) was established in 1905 in the USA. The first SAE meeting was held in 1906, and since then the Transactions of the Society have been published. In the USA, standardization work began in 1910 with the first issue of the *SAE Handbook*

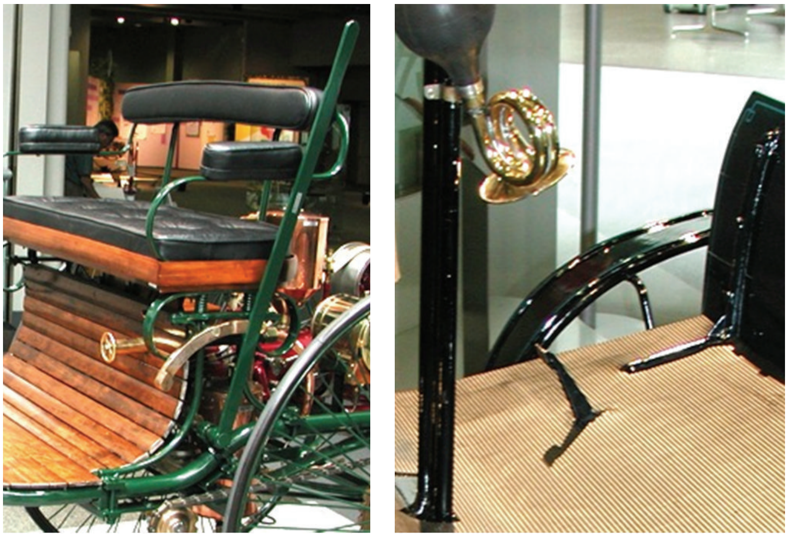

(a) (b)

Figure 2: Hand brake lever (Benz Patent Motor Vehicle 1886 (a)) and foot brake pedal (Benz Velo 1893 (b)) (the author's (MA) photo collection).


(a) (b) (c)

Figure 3: Meters on bulkhead (Alpha Romeo 24PH 1910 (a)), meters in instrument panel (Dodge Brothers Touring 1915 (b)), and meter cluster (Buick Series 50 1932 (c)) (the author's (MA) photo collection).

of Standards and Recommended Practices. The number of members reached more than 4,300 at the end of the 1910s [3].

In summary, the first human factors development was designing controls for the primary driving task, such as the steering wheel and the brake pedal, which allowed for operation of a heavy self-powered vehicle using only muscle power. The second development was introducing gauges to inform the driver about the mechanical condition of the vehicle and then driving condition (speedometer). In addition, industry associations established in this early stage, such as VDA and SAE, played important roles in the development and dissemination of information related to automotive technology.

2.2. The Dawn of Automotive Human Factors Design (1920–1939).

During these two decades, the basic controls and displays of the motor vehicle continued to evolve. An ignition-timing lever had accompanied the steering wheel from early on. Horn buttons began to be installed in the center of the steering wheel in the late 1920s.

With regard to information presentation, gauge clusters first appeared in 1920s, often on a separate panel. Grouping gauges allowed drivers to read them at a glance. However, most gauge clusters were placed in the center of the instrument panel.

Before the 1920s, switches or knobs typically did not include labels to indicate their function. Drivers had to learn and memorize the function of each. Labels first appeared on controls and on the surface of instrument panels in the 1920s.

In the 1930s, speedometers and other instruments began to be installed directly in front of drivers to improve their visibility (Figure 3(c)), a practice that became common in the 1940s. American and many European luxury automobiles in this period were equipped with a shift lever on the steering column.

In early vehicles, one signaled the intention to turn using a *winker*, a mechanically operated arm or flag that extended

Figure 4: Turn indicator (BMW 335 1939, courtesy of H. Bubb).

from the side of the vehicle, first appearing in the 1910s. The exterior signal became a mechanical semaphore in the 1930s (Figure 4) and, finally, an electric light in the 1950s in Germany. A turn-signal switch or turn-signal lever was also being installed in the steering column by the late 1930s (Figure 5).

The seat-sliding mechanism, which adjusts the driving position, appeared in the 1920s. It allowed drivers with different body sizes to find a reasonable distance between the pedals and the seat.

Until the 1930s, the focus of automotive technology was on meeting basic functional requirements, primarily mechanical, to provide a durable vehicle. The shift at that time was toward designing vehicles that could go faster, with the 1934 Chrysler Airflow and its emphasis on aerodynamics as an example. Consequently, cabs shrunk and the car body became more rounded. This, in turn led to efforts to design the cab layout to fit the human body size and provide increased seating comfort while maintaining outward visibility. In an early book on automotive engineering written by Wunibald Kamm, an automobile engineer and an aerodynamicist famous for his Kamm-tail theory, provided an example of desired cabin dimensions (Figure 6) [4].

Thus, basic human factors design features, such as easy-to-operate steering equipment and switches, visible gauges, and a reasonable driving position, were introduced during the 1930s and 1940s. Note that, throughout that period, design decisions to accommodate human operators and passengers were based largely on heuristics from engineers' experience. Also numerous features were designed and implemented to ease the driving task, such as synchronized gears and improved windshield wipers, as well as switchable low and high beams. For additional information on these and prior developments, see [1, 2, 5].

The number of traffic crashes increased after World War I as the production of automobiles increased. In 1920, German psychologist Narziss Ach outlined the importance of psychology and technology in preventing crashes from the perspective of a scientific discipline that he called *psychotechnik*, which is closely related to human factors [6]. At the end of the 1930s, Forbes pointed out that understanding the limitations of driver capabilities such as visual characteristics and reaction time, “human factors” in traffic crashes, was necessary [7]. Both engineers and psychologists were aware of the importance of the human element in vehicle design and traffic safety in this period.

3. Human Factors Activities after World War II until 1989: The Era of Occupant Accommodation and Safety

3.1. Establishment of Human Factors as a Field of Endeavor (1940 to 1949). Although one can identify the roots of human factors being in early work in industrial engineering, such as that of Taylor and Gilbreth, activities at Bell Labs on communication quality, and other examples, human factors as a profession did not take off until WWII [8]. Human factors research was introduced during World War II to adapt military technologies to human operators to make systems more effective and reliable [9–11]. This research field was then expanded to the commercial aviation and automotive industries after World War II.

There was not an immediate transfer of human factors ideas from military to civilian activities. In part, this was because the initial transfer was from military organizations to defense contractors, which took several years, and Europe and Japan were recovering from World War II.

However, this period was not without some progress. Passive-safety technology was introduced at the end of the 1940s. The instrument panel was covered with sponge rubber in American automobiles, by Tucker in 1948 and Chrysler in 1949.

Also, there was considerable growth in the organizations interested in automotive research, some shortly after World War II, others later. The earliest one was British Motor Industry Research Association (MIRA) (UK), founded in 1946.

The following sections briefly describe automotive human factors studies and their output (mainly standards) and outcomes (products) from 1950 to 1989 by decade. Table 1 summarizes the major developments for each decade.

3.2. Human Factors Research Activities in 1950s: First Decade of Human Factors Research. A survey of the literature on human engineering in the 1950s, conducted by the U.S. Army Human Engineering Laboratory [12], indicated that studies at that time focused on driving visibility (including glare), cab layout based on anthropometric data, and the design of controls.

With regard to anthropometry, in 1955, for the first time, the SAE published data that included 5th- and 95th-percentile values for use in cab layout (Figure 7) [13]. During this decade, research was also conducted on human-body

Table 1: Overview of history of human factors researches, their outputs and outcomes.

			Empirical human factors design					WWII						
			1886–1899	1900–1909	1910–1919	1920–1929	1930–1939	1940–1949	1950–1959	1960–1969	1970–1979	1980–1989	1990–1999	2000–2009
Primary driving task	Controls	Human factors research and output	Empirical: control by human muscle power			Empirical: access controls while holding steering	Empirical: access controls while holding steering		Symbols to indicate function of controls		Anthropometrical data for hand reach (SAE J287, 1976), Standardized direction of movement of control (SAE J1139, 1977)			
		Outcome	Steering wheel, Foot pedal,			Horn button and timing lever around steering wheel Labels indicate its function	Shift lever, turn signal lever around the steering column		Comprehension of function for people with different languages		Designing location of controls Common design of control direction	Steering wheel switch	Integrated joystick (Toyota, 1998)	i-Drive with controller nob in the center console (BMW, 2001), integrated center control (Nissan 2001)
	Displays	Human factors research and output		Empirical: obtain information about vehicle condition	Empirical: visible information	Empirical: decrease amount of eye movements to check meters and gauges	Empirical: smaller eye shift to access the meter cluster		Empirical: smaller eye shift to access the meter cluster Avoid influence of sunlight		Symbols for Motor Vehicle Control, indicators, and tell tales (SAE J1048, 1974, ISO2575)	Investigation of advantage of HUD for vehicle display (1970s)	Measurement of visual accommodation (Toyota 1998)	
		Outcome		Installing gauges and speedometer	Instrument panel with meters and gauges	Clustered meters in instrument panel	Meter cluster in front of driver		Meter cluster in high position with sunshade		Commonly used symbols	Introduction of HUD for vehicle display (GM, Nissan 1988)	Introduction of center on-dash meter (Toyota)	
	Visibility to road scene	Human factors research and output		Empirical: perceive approaching vehicles from behind	Empirical: road scene visibility in rain condition					Motor vehicle drivers' eye locations (SAE J941, 1965), Eyellipse (SAE Paper 680105, 1968) Passenger car rear vision (SAE J834, 1967)	Regulation for rear view mirrors (Directive 71/127/EEC, 1971), Field of view from automotive vehicles (SAE SP-381, 1973)	Backing sensor	Investigating visibility using digital human model Rear view monitor	Night Vision System
		Outcome		Rear view mirror	Windshield screen wiper					Define range of drivers' eye positions Examine direct visibility	Specifying visible area in rear view mirrors Investigating location of traffic signals, traffic signs, pedestrians in the driver's view		Time and cost effective design of visibility using CAD	

Table 1: Continued.

				Empirical human factors design				WWII						
			1886–1899	1900–1909	1910–1919	1920–1929	1930–1939	1940–1949	1950–1959	1960–1969	1970–1979	1980–1989	1990–1999	2000–2009
Driver Workspace	Seating and packaging	Human factors research and output				Empirical: adapt to women drivers	Design drawing		Anthropometrical data for human body (SAE SP142A, 1955), SAE Manikin Subcommittee (1959)	Defining and measuring H-point (SAE J826, 1962), 2DM, 3DM	Chrysler's Digital Human Model CYBERMAN (1974), Measurement of pressure distribution of seat	SAMMIE, CAD with digital human model, is in the market	Commercial digital human model, Ramisis, Jack Body movement analysis using motion capture system	Combining digital human model with CATIA and ALIAS CAD, estimating muscle load using DHM
		Outcome				Seat adjuster	Cabin space design			Cabin space and seat layout design Precise design of seating configuration based on H-point,	Designing seat back angle		Time and cost effective design of packaging	Evaluation of ingress/egress motion
	Vibration and comfort	Human factors research and output		Frederick Lanchester (UK) proposed the cabin movement is to be the same as that of human body while walking				Seat cushion and comfort (SAE 1940')	Motor Vehicle Seating Manual (SAE J782, 1954)		Relationship between mechanical vibration and discomfort (ISO 2631, 1974)	Analysis of resonance frequency of body parts	Evaluation of vibration discomfort in multiaxis environment (ISO 2631-1, 1997)	Temporal factor in vibration discomfort
		Outcome		Peak frequency of cabin vibration was about 2.0 Hz							Establish evaluation method for vibratory comfort	Designing cushion of seat	Evaluation to integrate vibration in multiaxis	
	Climate	Human factors research and output			Empirical: comfort in winter time							Thermal manikin	Equivalent temperature (SAE J2234, 1993)	Ergonomics of thermal environment of vehicle (ISO14505 series)
		Outcome			Cabin heater							Evaluation of vehicle climate	Evaluating thermal comfort using Equivalent Homogeneous Temperature	Evaluation of cabin thermal comfort, combining thermal manikin, calculation of EHT, and subjective evaluation

TABLE 1: Continued.

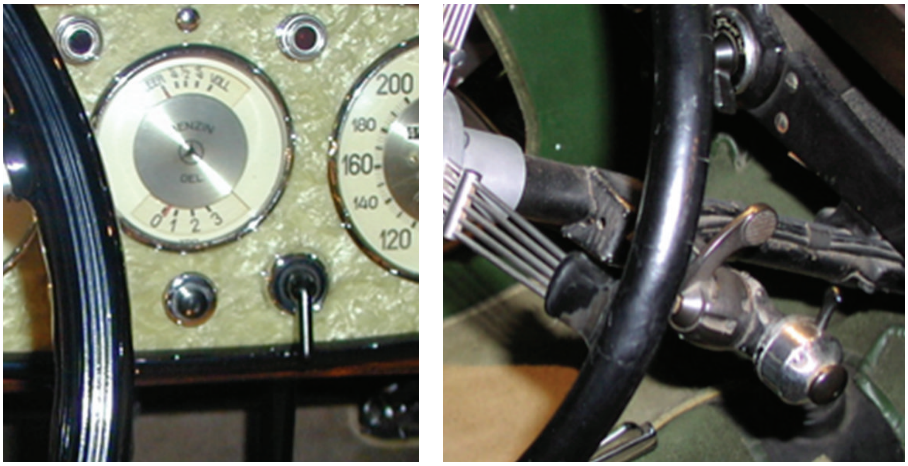
				Empirical human factors design				WWII						
			1886–1899	1900–1909	1910–1919	1920–1929	1930–1939	1940–1949	1950–1959	1960–1969	1970–1979	1980–1989	1990–1999	2000–2009
Driver's condition	Driver's Fatigue	Human factors research and output								Physiological measure for fatigue (HR, GSR, BP) CFF study Recommendation of having rest Evaluation of vehicle vibration		HRV as a measure of fatigue		
		Outcome												
	Impairment	Human factors research and output									Physiological measure for drowsiness, EEG, EOG, GSR	Partial eye closure as a measure of drowsiness (Dingus, 1986)	Percentage of eye closure as measure of drowsiness (PERCLOS)	
		Outcome										Algorism to detect drowsiness (NHTSA DOTHS808247, 1994)	Drowsiness detection system	
Crash injury		Human factors research and output						Instrument panel covered by sponge rubber (Tucker 1948, Chrysler 1949)	Investigation of body damage by accident Crash test using high speed camera Crash dummy (GM, Ford)	Crash dummy, FMVSS208 frontal crash test in 30 mph (1966)	Crash dummy, Hybrid II (1974), Hybrid III (1978) Injury index AIS-1971, AIS-1976	Dummy, Euro-SID-1 for side impact (1989)	Dummy for side impact, Bio-SID, more sensors and more biofidelity, (1990) and SID-II (1994)	
		Outcome							Seat belt (Nash, 1949) Nondeformable passenger cell (Daimler-Benz 1952) 3-point seatbelt (Volvo, 1959) Head restraint (AM, 1959)	Collapsible steering column (GM, 1967) Mandatory belt use in front seat (1967, USA; 1969, Japan)	Standardized assessment method	Airbag	Side impact bar, Side airbag	
Interaction with driver information system		Human factors research and output								There was tolerable visual sampling duration while driving (Senders, 1967)	Measurement of eye movement on in-car displays (Mourant, 1978)	Visual sampling model and glance time study (Wierwille)	Visual behavior and driving performance as a measure of distraction (Zwahlen, 1986), mental workload measures	Workload measurement for distraction: occlusion method (ISO16673, 2007) and LCT (ISO26002, 2011)
		Outcome									Single glance time was 0.5–1.5 seconds. Several glances for radio task	Assessment of IVIS using glance time	Discussion of map display JAMA guideline ver. 1.0 (1990), ver. 2.0 (1999) HARDIE guideline (1995)	Guidelines (JAMA ver. 3.0 (2004), AAM etc.) for information device

Table 1: Continued.

		Empirical human factors design					WWII						
		1886–1899	1900–1909	1910–1919	1920–1929	1930–1939	1940–1949	1950–1959	1960–1969	1970–1979	1980–1989	1990–1999	2000–2009
Interaction with advanced driver-assistance system	Human factors research and output												Overreliance, overtrust, and situation awareness with ADAS Analysis of driving behavior
	Outcome							Conventional cruise control			Antilock Braking System (ABS)	Electronic stability control (ESC), vehicle stability control (VSC) Adaptive cruise control (ACC) in the market, night vision	Lane-keep assist Full-range ACC Collision mitigation braking system
External communication access	Human factors research and output								Visual demand while driving (Senders, 1967)			Accident statistics analysis for risk of mobile phone use while driving	Naturalistic driving study
	Outcome									Cellular automobile phone service (ARP Finland, 1971; NTT, Japan, 1979)	Analog cellular service (US 1984)	Prohibition of hand held use of cellular phone (Swiss, 1996; Japan, 1999)	Regulations for use of cellular phone (USA, 2001 (NY), Germany, 2001; France, 2003; UK, 2003)
Driving behavior	Human factors research and output								Driver vehicle control model Development of early DS (GM)	Visual behavior while driving using eye tracker	Researches on UFOV	Visual attention measured by peripheral detection task	Naturalistic Driving study Driving behavior study using DS
Related technologies and vehicles' environment		Very rough road	Increased speed		Radio tuner was installed		Highway mobile telephone (USA, 1946)		Development of highway		RDS-TMC (EU)	Telematics service (US OnStar 1995, Germany TeleAid 1997 BMW Assist 1999, Japan MONET 1997, Carwings 1998)	Smart phones

Table 1: Continued.

		Empirical human factors design				WWII						
	1886–1899	1900–1909	1910–1919	1920–1929	1930–1939	1940–1949	1950–1959	1960–1969	1970–1979	1980–1989	1990–1999	2000–2009
Organization and academic society		VDMI (currently VDA, Germany, 1901), SAE (USA, 1905)		SIA (France, 1927)		JSAE (Japan, 1947), FISTA (1948)	HEFS (USA, 1956), IEA (1959)	JSAE automotive ergonomics study group (Japan, 1962)		HFES Europe Chapter (1983)		
Public institutions				VTI (Sweden, 1923)	TRL (UK, 1933)	MIRA (UK, 1946), TNO human factors (The Netherland 1949)	TTI (USA, 1950), BASt (Germany, 1951)	ONSER (road safety org France, 1961), UMTRI (USA, 1965), JARI (Japan, 1969), TNO Traffic Behavior Department (The Netherland 1969)	NHTSA (USA, 1970), HUSAT (UK, 1970), IRT (France 1972)	INRETS/LESCO (combining ONSER and IRT, currently INFST-TAR/LESCOT) (France 1986)		
Conference meetings		SAE conference (USA, 1906)		TRB (USA, 1920)		FISITA congress (1947)	JSAE conference (Japan, 1951), HFES meeting (1957)	IEA congress (1961), Stapp (USA, 1962)	ESV conference (1971)	Vision in Vehicle (1985)	AVEC (1992), Driving Simulator Conference (1994), ITS World Congress (1994)	Driver Assessment Conference (USA, 2001), International Conference on Driver Distraction and Inattention (EU, 2009), AutomotiveUI (2009), HUMANIST (EU, 2008)
Social background		Speed violation penalty		Increase of number of vehicles (USA)			Increase of number of vehicles (Europe)	Media promotion for safety (Chevrolet, Corvair USA)	Media promotion for safety (Ford Pinto, USA)	Media promotion for safety (Jeep CJ05, Suzuki Samurai, Audi 5000, USA)	Media promotion for safety (GM CK pickup, USA)	Media promotion for safety (Ford Ram, Crown Victorias, Explorer, USA)



(a) (b)

Figure 5: Turn signal lever in instrument pane (Mercedes-Benz 500K 1935) and that in steering column (Morris Eight Series I 1937) (the author's (MA) photo collection).

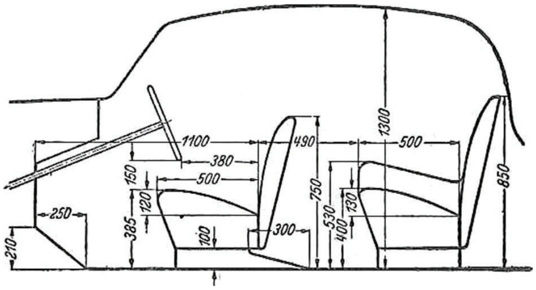


Figure 6: Cabin dimensions shown in Kamm's book "*Das Kraftfahrzeug*" (1936).

injuries caused by vehicle crashes [14]. Experimental technologies for crash tests (e.g., dummies, accelerometers, and high-speed cameras) were developed [15].

Following up on some advances in passive safety earlier in the 1940s, Nash Motors installed the first seatbelt in 1949. Other American manufacturers introduced seatbelts in the 1950s. In 1952, Barényi, an engineer at Daimler Benz, invented the nondeformable passenger cell and in later years, the crumple zone and collapsible steering column.

Some European vehicle manufacturers in this period introduced symbols to indicate the functions of controls. The position of the gauge cluster was raised to be closer to the normal line of sight and, therefore, was easier to read.

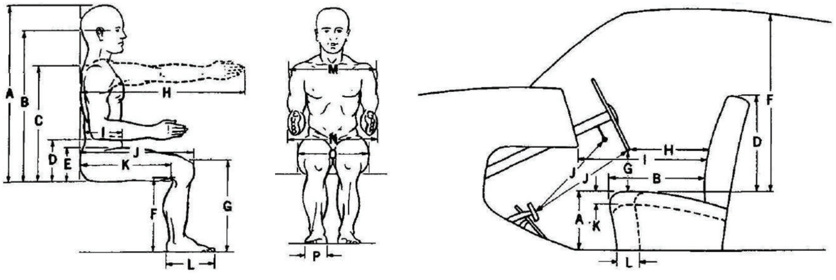
Subsequent to MIRA's founding in the UK in 1946 was the founding of Texas Transportation Institute (TTI) (US, 1950), German Federal Highway Research Institute (BASt) (Germany, 1951). Also established around this time were organizations specifically focusing on safety and human factors—TNO Human Factors (The Netherlands, 1949), ONSER (road safety organization, currently INFSTTAR, France, 1961) and the automotive ergonomics study group in JSAE (Japan, 1962).

A variety of automotive human factors research efforts began during this period. Methods from psychology, medicine, and anthropology were introduced. An important method involved using statistical distributions of anthropometric dimensions to establish vehicle design standards for those dimensions. This method directly linked human factors research to production of vehicles geometrically adapted to human characteristics, a method that was developed further in the next decade.

3.3. Human Factors Research Activity in 1960s: The Decade of Anthropometry. In the 1950s, automobile manufacturers recognized that anthropometric data could be the basis for laying out the cab to ensure that the driver (1) could see the road, traffic signals, and other vehicles outside of the cab, (2) could see controls and displays inside the cab, and (3) would be able to reach controls. In 1959, the SAE Manikin Subcommittee began developing an easy-to-use tool for ergonomic design based on anthropometric data. The SAE two-dimensional manikin (2DM) and three-dimensional manikin (3DM) were codified in SAE J826, which was published in 1962 [16]. The hip-point (H-point), which was the origin on the human body for automotive cab design, was defined in this standard together with specific measurement procedures. The 2DM was used to design the side view of the vehicle, and the 3DM was used to design cab mockups.

Based on the anthropometric research, the driver's eye position was defined in SAE J941, and the concept of the eyellipse, which specified the range of the driver's eye position, was developed [17–19]. What drivers of widely varying body sizes would be able to see could be examined using the eyellipse. Standards for front-view and rear-view visibility were also published (SAE J834, 1967) [20].

At that time, automobiles were commonly used in the USA and driven by a wide range of people. Therefore, the US car manufacturers were motivated to collect anthropometric data for cab design to accommodate that range



(a) (b) (c)

Figure 7: Human body measurements and vehicle dimensions shown in SAE SP142 (1955).

Figure 8: Field experiment of Critical Fusion of Flicker (CFF) measurement for highway drive in Japan (Brochure of IPRI, AIST, 1969).

of drivers [21]. This data was also helpful to car manufacturers outside the USA who were developing cars for export to the USA and served to further improve various SAE standards that had been developed or were in development.

Frontal-crash test procedures to protect occupants were introduced in FMVSS 208 [22]. In 1959, Volvo was the first manufacturer to provide three-point seatbelts. In the same year, American Motors also equipped their automobiles with head restraints to avoid neck injury in rear-end collisions. In 1967, General Motors conducted pioneering work on collapsible steering columns designed to reduce chest impact injuries [5].

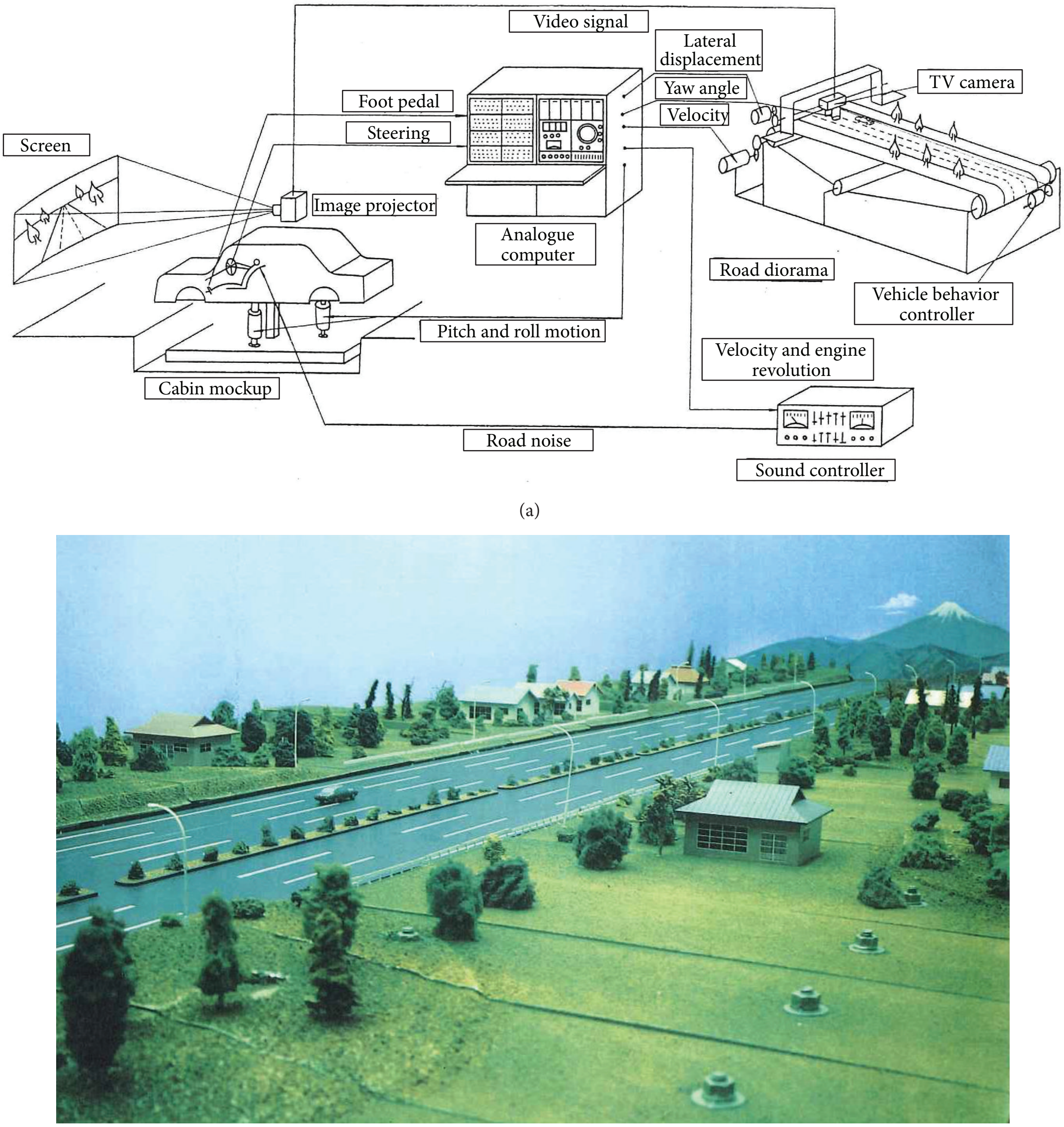
The construction of special-purpose, high-speed roads began with the first autobahn in Germany in the 1930s, followed by construction of interstates (USA), autoroutes (France), motorways (UK), and autostrada (Italy) beginning in the 1950s, and followed by significant highway construction in Japan in the 1960s. Because trips on such roads tended to be long, driver fatigue became a concern. There were many studies done in Japan, mainly by researchers with medical backgrounds, to evaluate driver fatigue using such physiological variables as heart rate, GSR (galvanic skin response), blood pressure [23], and CFF (critical frequency fusion) (Figure 8).

Figure 9: Helmet for occlusion device (courtesy of J. W. Senders).

With the development of control theory, studies were conducted to apply this theory to steering maneuvers [24–28]. Studies to measure mental workload, introducing methods from physiology and the cognitive sciences, began in the 1960s. Brown and Poulton assessed drivers' spare mental capacity using auditory subsidiary tasks requiring the driver to identify a digit that differed from the previous one [29]. One pioneering study on driving behavior was Sender's 1967 study to measure visual demand while driving, using an occlusion device with a moving frosted plastic visor on the helmet (Figure 9) [30].

During the 1960s, driving simulators were developed to study vehicle dynamics and to analyze driving behavior. It is not certain when the first simulator was developed, but there were driving simulators in the 1950s. General Motors developed a driving simulator using a gimbal structure to give pitch and roll motion to the driver [31]. The driving simulator developed in 1976 by the Mechanical Engineering Laboratory of AIST (Japan) had a moving cab, and the driving scene was obtained through a movie camera running a miniature diorama of a road in town and in a rural area (Figure 10) [32]. Driving simulators were also developed in US universities. Interestingly, it was not until about 17 years later, with the advent of the Daimler-Benz simulator, that driving simulators received broad attention [33].

In the USA, a major factor in the movement to improve crash safety was the investigative news media. The first vehicle to attract attention was the 1961–1963 Chevrolet Corvair, which in a sharp turn, had a tendency to spin and/or rollover. The Corvair was an unusual rear-engine vehicle, and there



(a)

(b)

Figure 10: Driving Simulator of Mechanical Engineering Laboratory of AIST (1968) (Technical Report of MEL, no. 89, 1976).

was considerable discussion of its suspension system in a book by Ralph Nader, a consumer advocate [34]. The book's title, *Unsafe at Any Speed,* captured the way some felt about Corvairs. As a result, there were congressional hearings about vehicle safety (that led to a black eye for General Motors), eventual withdrawal of the Corvair from production, and a significant increase in interest in vehicle safety.

The interest in safety led to the establishment of the Highway Safety Research Institute at the University of Michigan, now the University of Michigan Transportation Research Institute (UMTRI), in 1965 and the National Highway Traffic Safety Administration (NHTSA) in the U.S. Department of Transportation in 1970. TNO in The Netherlands started a Traffic Behavior Department in 1969, which focused on traffic safety. In the same year, Japan Automobile Research Institute (JARI) was founded. They joined a worldwide collection of organizations (see Table 1).

The growth in the worldwide production of automobiles led to increased interest in designing vehicle cabins suitable for a wide range of people. As the number of traffic accidents rapidly increased with increased production, safety became a major concern for society. Automotive safety technology had evolved since the last decade, but it was facilitated by news media in this decade. Human factors research led first to advances in passive safety and later to advances in active safety. Research on measurement of fatigue, mental workload, and driving-task demand developed in this decade. A shift in human factors research began from a focus on physical characteristics to cognitive characteristics.

3.4. Human Factors Research in the 1970s: Establishing Crash-Safety Assessment and Occupant Comfort. The impact of the US news media in bringing attention to crash safety continued in the early 1970s, focusing on the Ford Pinto. When struck from the rear under certain circumstances, Pintos would dramatically catch fire [35–37], videos of which are still available (http://www.youtube.com/watch?v=rcNeorjXMrE, http://www.youtube.com/watch?v=lgOxWPGsJNY). A critical document in the case was a cost-benefit analysis done by Ford, which compared the cost of making changes to the vehicle to prevent or reduce fires with the cost of injuries and lives lost, an idea that has been the source of numerous ethics discussions over time. However one feels about the Pinto, the case generated an intense focus on vehicle safety, in particular with regard to fires and safety in crashes, especially rear-end crashes. As with the Corvair, the Pinto's poor publicity led to a sharp decline in sales and eventual withdrawal of the Pinto from production. The Pinto case served as the stimulus for further research in the USA.

To help prevent rear-end crashes, Irving and Rutley investigated staged signaling concepts for different braking levels, conveying more information to following vehicles, concepts that led to improved braking over those in use [38]. Also the number and position of brake lights varied, leading to the idea of center, high-mounted stoplights. The effectiveness of the high-mounted stoplight was studied in the 1980s [39, 40]. During this decade, there were also studies of nighttime visibility distance of different headlight beam patterns and technologies (conventional tungsten, sealed beam, and halogen), as well as their effects on glare [41].

Improved understanding of what happened in crashes was also a research focus. Crash dummies were developed by several different organizations. They were integrated into Hybrid I in 1971 and Hybrid II in 1974. Sensors in Hybrid II were located in the head, chest, and femur. To make the dummy more realistic, Hybrid III was developed in 1976 [42]. Ten sensors were located in the head, neck, upper body, femur, knee, and leg, where injury might occur in the event of a crash. The severity of injury of each part of the body could be assessed based on the acceleration of each location. Head Injury Criteria (HIC) were defined by NHTSA in 1971 to assess the severity of head injury using the dummy. The Abbreviated Injury Scales (AIS-1971 and AIS-1976) for determining the level of injury produced by actual accidents were also established during this decade. The assessment method was standardized during this period [43].

However, crash safety was not the only topic of interest during the 1970s. Based on anthropometric research, an SAE standard for hand reach was published in 1976 (SAE J287) [44, 45]. To reduce driver confusion when operating controls, the direction of the movement of controls was standardized in SAE J1139 in 1977 [46].

Symbols to indicate control functions were introduced in the 1950s, mainly for European cars, to avoid the need to produce a different instrument panel for each language region in which a vehicle was sold. These symbols did not require reading written words and were intended to be intuitive. However, when different symbols were used to indicate the same function, drivers could become confused. To avoid such confusion, standard SAE J1048 was established in 1974 [47].

Studies on vehicle vibration and comfort have been conducted since the 1940s. Vibration and shock may cause low back pain and performance changes [48]. Vibration of the vehicle's cab occurs along all three axes, both linearly and rotationally. The most important is vertical movement transferred though the vehicle suspension and car seat. A method to estimate the perception of discomfort was standardized in ISO 2631 in 1974 [49].

In addition to specific research topics, research tools were developed and improved in this decade. Eye trackers, devices used to measure eye-gaze location, became available for vehicle and simulator use in the 1970s. For example, Mourant, Rockwell, and others measured glance time to the mirrors, radio, and the road while driving for novice and experienced drivers [50].

The driving simulator became a tool in human factors research. Volkswagen developed a driving simulator with a three-axis gimbal. A CRT display was used to present a road scene that involved a computer-generated line drawing. Various sounds were also presented. This driving simulator was used to investigate the driver's evasive behavior [51]. A driving simulator using a linear rail was developed at Virginia Tech in 1975 [52].

This most noteworthy result of this decade was the translation of human factors research into practice. Various standards were prepared to design controls and to evaluate seating comfort. Crash dummies were established and utilized by the New Car Assessment Program (NCAP), which began in 1979 in the USA.

3.5. Human Factors Research in the 1980s: Computer-Aided Design for Automobiles, Cab Comfort, Rollovers, and Assessment Methods. As with every recent decade in the USA, the 1980s had a particular vehicle that received attention for issues related to crashworthiness. That vehicle was the Jeep CJ-5, whose rollover propensity was the subject of a broadcast by *60 Minutes*, the most-watched investigative news program on US television. The critical episode, broadcast on December 21, 1980, showed Jeep CJ-5s rolling over when making sharp turns. What many fail to recall is that there was supporting statistical data showing that the CJ-5 was much more likely to roll over than other similar vehicles [53, 54]. For an interesting summary, see [55]. The CJ-5 problems served to spark human factors research on vehicle handling.

Another vehicle that received attention in that decade was the Suzuki Samurai, a short wheelbase, four-wheel drive utility vehicle with a propensity to roll over. Suzuki had a very bitter legal battle with the Consumer's Union, which publishes the most popular consumer magazine in the USA, *Consumer Reports.* Unusually, the vehicle was rated as "not acceptable." Sales dropped from 77,500 vehicles in one year to 1,400 the next year. Suzuki sued the Consumers Union but lost, and the production of the Samurai ceased. The Suzuki case emboldened safety advocates who had been sometimes reluctant to challenge the auto companies with "deep pockets" to fund protracted legal actions.

Allegations of unintended acceleration of the Audi 5000 were publicized on *60 Minutes* on November 23, 1986 [56]. Again, given the bad publicity, sales of the Audi 5000 plummeted from 74,000 vehicles in 1984 to 12,000 in 1991. Ironically, the final verdict from the U.S. Department of Transportation was that, while there were design aspects that could startle drivers or contribute to a higher incidence of pedal misapplication, there was nothing requiring a defect notification [57]. The important point here is that this is probably the first time that questions raised by the news media about vehicle safety were not supported by further investigations.

Interestingly, in recent years, there again have been questions raised concerning unintended acceleration; this time was for Toyota vehicles. *Dateline NBC* was the program involved, but in some ways the Toyota case is strikingly similar to that of the Audi 5000. There were allegations of trapped floor mats and concerns about failure of the electronic control systems, a claim that was debunked by NASA [58]. Again, Toyota sales suffered as a consequence, but no vehicles were withdrawn from the market.

In 1980, Brown stated that the improvement in the crash statistics "has undoubtedly resulted from technological advances in the design of steering, braking, tires and suspension systems, affording the driver better control of his vehicle" [59, pages 3–14]. He also emphasized the importance of optimizing information presentation in the vehicle and introducing objective evaluation and quantification instead of pure subjective assessment.

New tools for designing cab dimensions and visibility were developed in the previous decade. Chrysler developed CYBERMAN, a digital human model (manikin) in 1974. However, it was simple and its usefulness was limited. The System for Aiding Man-Machine Interaction Evaluation (SAMMIE) was developed in the UK for a consulting service for ergonomic design by SAMMIE CAD, Ltd., in the 1970s. The three-dimensional, digital human model consisted of 21 links and 17 joints. The cab dimensions and layout of controls in the cab could be evaluated by specifying the joint angles of the three-dimensional human model based upon anthropometric data of representative drivers. Various digital human models were developed during this period. Linked with computer-aided design (CAD), digital human models worked effectively. SAMMIE worked with SAMMIE CAD system, but interchangeability with other systems was limited. Jack (USA), RAMSIS (Germany), and other digital human models were developed during this period. RAMSIS could link with the CATIA CAD system, which was and still is the most commonly used CAD system in the automotive industry. Compared with the traditional anthropometric data and hard manikins, digital human models can lead to shorter development times of vehicle cabs, reduce development cost, and lead to cabs that accommodate a larger fraction of the population [60, 61].

Head-up displays (HUDs) were initially developed for aviation and superimpose information of aircraft air speed, altitude, and angle of attack onto the forward view. As eye transition and accommodation times were reduced, the user could spend more time looking at the forward scene. In motor vehicles, HUDs have been used to show vehicle speed, warnings, turn signals, and more recently, navigation information. The first studies with HUD prototypes were conducted by Rutley [62], who showed that HUDs can have benefits without the negative distracting effects reported in aviation applications [63]. HUDs were introduced in the market at the end of the 1980s (General Motors 1988, Nissan 1988). As the initial application was to present speed, which was not as time-critical as the flight data shown in aircraft, the customer demand for automotive HUDs when introduced was not great.

Also occurring at this time was considerable research to assess human thermal comfort [64], research that has its origins in Willis Carrier's development of the psychrometric chart [65]. The factors contributing to human thermal comfort, air temperature, radiant temperature, air velocity, humidity, metabolic rate, and the distribution and insulating value of clothing were not all easy to measure in a real vehicle cab. To evaluate space-suit thermal comfort, in 1966, NASA developed a thermal manikin that had a three-dimensional human body and simulated the heat transfer between the human body and the thermal environment. By the end of the 1970s, thermal manikins were used to estimate thermal comfort in vehicle cabs [66].

Drowsiness while driving increases crash risk. A driver's drowsiness, arousal level, and fatigue can be measured using such physiological variables as EEG (electroencephalogram), heart rate, respiration rate, and GSR (galvanic skin response) [67]. As was shown in early studies, physiological measures could be reliably measured in experimental conditions and provided useful information. However, it was difficult to convert the research into practice and develop a commercial drowsiness-detection system, primarily because wired sensors were needed. Thus, in the 1980s there was a shift towards noncontact image sensors (video cameras) that looked for slow eyelid closure to detect drowsiness [68]. Studies were conducted to obtain quantitative measures based on video images, and in the next decade PERCLOS (percentage of eyelid closure time) was established as the index of drowsiness [69]. In 2008, Toyota introduced a crash-mitigation system with eye monitor that detected eyelid closure and warned the driver.

Workload-measurement methods were established during the 1970s [70]. These methods used subjective measures (the Cooper-Harper scale, SWAT-the Subjective Workload Assessment Technique, and NASA TLX-the Task Loading Index), primary task performance measures, secondary task measures (from the task loading and subsidiary task methods), and physiological measures (EEG, pupillary response, eye movement, and heart-rate variability). They were used to measure mental workload while driving. Miura collected detection-reaction times to the illumination of small bulbs located around the front window, as the subsidiary task, to measure the useful field of view [71]. Results indicated that the useful field of view became smaller, and the reaction time of a detection task became longer as task demands increased (e.g., driving in crowded traffic).

With increasing computer power, large driving simulators were developed in the 1980s. In the 1970s, VTI of Sweden began developing a driving simulator with a two-axis gimbal and a linear rail. It had a wide screen and was controlled by a detailed vehicle-dynamics model [72]. An example of its use, which began in 1983, was the investigation of driving on slippery roads and the effects of alcohol. The major development was the Daimler-Benz high-fidelity driving simulator with a motion system that combined the hexapod motion platform and two-dimensional linear rails. A full-size vehicle was placed in the dome on the motion platform. It was introduced in 1984 and was used to investigate active-safety systems, vehicle dynamics, and other topics [73]. During the 1980s, various driving simulators were developed in the USA, Europe, and Japan [74]. Common topics in the 1990s included studying driving behavior in risky conditions, the use of driver information systems [75–78] and the use of advanced driver-assistance systems (ADAS) [79, 80] and the effectiveness of warning systems of various types. One example was using the pedals and steering wheel to provide active feedback to facilitate drivers' performance of a recommended action [81].

The end of the 1980s saw the beginning of an era of driver information and driver-assistance systems (see the next section). One early human factors study of driver information systems involved measuring glance time and number of glances for a variety of conventional tasks and navigation tasks using a prototype computer map navigation system [82]. One study indicated that centerline deviation increased when the driver used a CRT touch screen [83].

The 1980s were the decade of the computer. Digital computers and software began to see wide use in human factors research, including digital human models for designing cabin accommodations, thermal manikins for evaluating thermal comfort in the cabin, and video systems for measuring drowsiness. Computer technology reduced design time and made handling complex data easier. The questionnaire and the secondary-task methods were established for mental-workload measurement based on resource models from psychology. These measurement methods and driving-simulator technology would become useful human factors research tools for the intelligent vehicles and connected vehicles in the following decades.

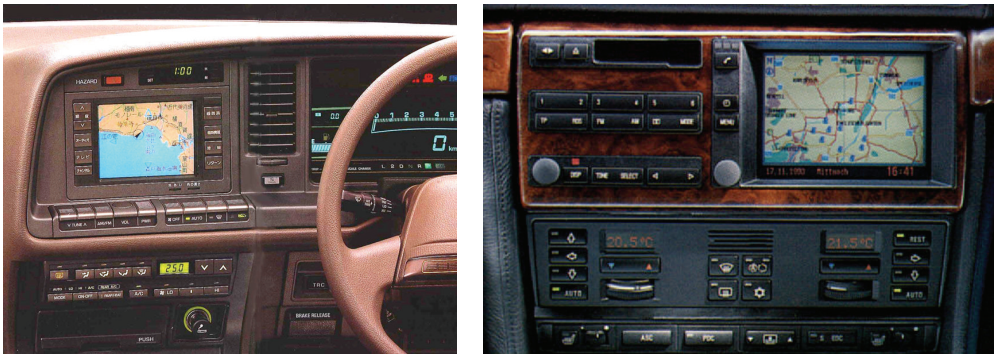
4. Human Factors Research Since 1990s: The Era of Intelligent Vehicles and Connected Vehicles

4.1. Driver Information Systems and Driver Distraction. Research on automotive human factors reached a turning point in 1990 with the introduction of Intelligent Transportation Systems (ITS), previously known as Intelligent Vehicle Highway Systems (IVHS). With the aim of enhancing vehicle mobility and safety using information and communication technologies, government projects began in the USA and Japan. The Electronic Route Guidance System (ERGS) was conducted in the late 1960s in the USA [84]. The Japanese projects included the Comprehensive Automobile Traffic Control System (CACS) (1973), Road/Automobile Communication System (RACS) (1984), Advanced Road Transportation System (ARTS) (1989), and Vehicle Information Control System (VICS) (1990) [85]. Europe's research initiative Programme for European Traffic of Highest Efficiency and Unprecedented Safety (PROMETHEUS) (1987–1995) initiated the research era of driver information and driver-assistance systems [86]. PROMETHEUS was followed by a sequence of projects (e.g., DRIVE, GIDS, EMMIS, HASTE, and AIDE) that focused on the development of integrated HMI concepts [87] and suitable evaluation methods [88].

The automotive industry also promoted ITS technology developments during this period. In 1981, Honda released Gyrocator, the first in-vehicle navigation system with a map using a transparency sheet. At about the same time, Toyota released NAVICOM, which indicated the direction of a destination using a simple arrow. Etak Navigator, the first after-market car navigation system using a digital map, was released in 1985 in the USA. The digital map was stored in cassette tapes and location was determined by a dead-reckoning system using a compass. In 1987, Toyota launched Electro Multi Vision, which was a predecessor of present-day, in-vehicle car navigation systems (Figure 11). An in-vehicle navigation system manufactured by Sumitomo Electric was installed in the Nissan Cima in 1989 [89]. The Bosch Travelpilot was delivered in the same year in Europe. In-vehicle navigation systems spread after GPS became available in 1990 (officially in 1993). The first on-board installed navigation system including a GPS unit and map material in Europe was delivered in 1994 by BMW using a color-TFT display and a button-operated software menu system. Later versions, which supported audio and communication functions, were moved to the center console and/or operated by a touchscreen, depending on the OEMs human-machine interface (HMI) concept. This development steadily led to unique integrated solutions for each brand as well as unique mobile navigation systems.

There were various efforts to design integrated driver interfaces for in-vehicle information and other existing in-vehicle systems (audio and climate) as the number of functions was increased. Toyota developed the integrated joystick (Toyota Ardeo 1998). BMW introduced i-Drive. Mercedes introduced Command. Audi introduced MMI (Multi Media Interface) as well (2001), which similarly included interaction using a rotary control knob in the center console [63]. Nissan introduced its integrated driver interface in the same year (Figure 12). The position of a central information display close to the windscreen became common at the end of the 1900s.

As with other decades in the USA, the 90s was not without its media controversies over crash risk, the most noteworthy of which was the 1977–1983 CK pickup, the most popular vehicle sold by General Motors. In a very dramatic



(a) (b)

Figure 11: Early car navigation systems (Toyota 1987 (a) and BMW 1994 (b)).

Toyota Ardeo 1998 Nissan Primera 2001 BMW i-drive 2001

(a) (b) (c)

Figure 12: Controls for in-vehicle information systems.

presentation on NBC's *Dateline* in 1993, a very popular news investigative show, a CK was shown being struck in the side and bursting into flames. The allegation was that the fuel tanks, mounted outside the frame rails, were vulnerable and could lead to fires if struck. Interestingly, careful investigation by General Motors found that the crashes had been staged, and rocket igniters had started the fires. In response, NBC retracted the story and paid General Motors for the cost of their investigation [90]. This was a huge blow to the news media and reduced its influence in advocating for auto safety.

Until 1990, the driver was regarded as an element of the driver-vehicle system, interacting with the vehicle by operating the steering wheel and pedals to manage the primary driving task. When a navigation system was installed inside the vehicle, the driver had to perform not only vehicle-control tasks by operating the vehicle, but also navigation tasks. When drivers used a paper map, reading the map while the vehicle was in motion was not easy. Often drivers had to stop the vehicle and read a map to find their way to a destination. When a navigation system was installed inside the vehicle, and the system indicated where to turn, the navigation task could easily be performed in parallel with driving tasks (i.e., a dual-task condition).

Concerns about excessive task demands led to studies of mental workload, human cognitive activity, and what is now commonly known as driver distraction. The 1990s saw the delivery of such guidelines as JAMA Guidelines (version 1.0 in 1990, and version 2.0 in 1999), UMTRI Guidelines (1993) [91], TRL Checklist (1999) [92], HARDIE Guidelines (1996) [93, 94], German Code of Practice [95] and other guidelines in Europe [96]. They gave descriptive principles for designing in-vehicle information systems. Also, relevant ISO activities were initiated to develop standardized evaluation methods and formulate minimum standards for in-vehicle HMIs [97].

Studies by Wierwille et al. and Zwahlen et al. in the previous decade suggested that glancing behavior could be an objective measure of driver distraction [98, 99]. Eye-glance evaluations are most readily conducted for information systems that have been developed and are available for on-the-road use. However, the systems must be assessed during the development. The occlusion device developed by Senders in the 1960s (see Figure 9) to measure visual demand in driving was used to simulate glance behavior during driving [100, 101]. Studies using the occlusion method with liquid-crystal shutter goggles were conducted under the aegis of the Alliance of Automotive Manufacturers (AAM) (USA),

Figure 13: Occlusion method with the shutter goggles.

ISO/TC22/SC13/WG8, and the Japanese Automobile Manufacturers Association (JAMA) to assess the level of distraction caused by visual-manual tasks and their degree of interruptibility (Figure 13) [102]. This method was internationally standardized as ISO 16673 in 2007 [103] based on input of the Advanced Driver Attention Metrics (ADAM) and Adaptive Integrated Driver-vehicle interface (AIDE) projects among others. In 2004, JAMA delivered JAMA Guideline version 3.0, which prohibited tasks that required a total glance time of more than 8 seconds [102]. SAE Recommended Practices J2364 (15-Second Rule and another occlusion procedure) [104], SAE J2365 [105] (task time estimation), and other procedures were also published as a result. In the search for entry methods that were less demanding than visual-manual interfaces, speech interfaces were examined [106, 107].

Several international design guidelines for in-vehicle information systems have been developed mainly in ISO/TC22/SC13/WG8 since 1994. Published standards were ISO 15005 (dialogue management) and ISO 15007 (measurement of visual behavior) in 2002, ISO 15008 (visual presentation) in 2003, ISO 17287 (suitability of TICS while driving) in 2003, ISO 15006 (auditory information) and ISO TR 16951 (criteria for determining priority of messages) in 2004 [108–112]. ISO 26002 (simulated lane change test, LCT) was published in 2011 for assessing driver distraction based on research from the ADAM project in Europe [113]. LCT was developed to evaluate visual manual secondary tasks but also cognitive loading tasks that used speech interfaces or involved phone conversations [114]. Burns et al. give an overview of the relevant evaluation methods [115].

Driver information systems have been developed as research projects since the 1970s and yielded commercial products such as car navigation systems in the 1990s. During their development, researchers were aware of the importance of human factors because using driver information systems while driving was quite different from using conventional in-vehicle equipment, with some ideas from studies of human-computer interaction for office work providing useful insights. In contrast to conventional in-vehicle systems, drivers could be confronted with a large amount of real-time information with which they interacted while driving. Measurement techniques for mental workload, glancing/visual behavior, and task demand developed in the last decade were applied to assess the amount of effort to use these driver information systems. Human factors researchers also played important roles in establishing guidelines and standards that offer principles for designing the systems in advance and evaluation methods accompanying the development process. Having guidelines and standards that were publicized by common agreement facilitated entrenching this technology in society.

4.2. Human Factors Research for Advanced Driver-Assistance Systems.

The 2000s were another decade in which crash safety received attention in the news media in the USA. High-profile media stories included (1) rollovers of Ford 15-passenger vans (picked up by several television programs), (2) rear-end crashes and subsequent fires involving 2005–2007 Ford Crown Victorias (commonly used as police cars), picked up by both NBC *Dateline* and CBS, and (3) rollovers of the 1998–2001 Ford Explorer. The Explorer received the most attention, including a segment on 60 Minutes and an entire hour on the PBS Frontline program (http://www.pbs.org/wgbh/pages/frontline/shows/rollover/etc/script.html, http://www.pbs.org/wgbh/pages/frontline/shows/rollover/etc/video.html). The Explorer problem was a combination of a high center of gravity combined with a narrow track width, along with failures of particular Firestone tires, which resulted in rollovers [116, 117]. One of the consequences of this matter was the passage of the US government's Transportation Recall Enhancement, Accountability and Documentation (TREAD) Act, which led to new tire-labeling standards, requirements for tire-pressure monitoring systems, and other changes.

Automobile safety technology began with efforts to reduce the consequences of crashes, by designing vehicles that would be less lethal when struck. Over time, there has been somewhat of a shift in human factors research towards active safety, seeking ways to prevent crashes.

The antilock braking system (ABS), first introduced in 1970, marked the formal beginning of active-safety technology. In 1990, electronic stability control (ESC) and vehicle stability control (VSC) came into widespread use. Adaptive cruise control (ACC) systems, which allow a vehicle to follow the preceding vehicle automatically by maintaining a preset time gap, were introduced by the end of the 1990s [118–120].

In addition, backup monitors utilizing the navigation system's display were introduced in the 1990s to reduce backing crashes. The 2000s saw the introduction of lane-keeping systems, which assist drivers by steering to help them stay in the lane (Nissan 2001), and the collision-mitigation braking systems, which intervene with active braking when distance-sensor data indicates that a collision is unavoidable. These systems are an extension of lane-departure warning systems and blind-spot warning systems. Recent entries into the market are the lane-change decision-aid systems, which provide warnings when the driver begins to change lanes, but another vehicle is in the adjacent lane.

Projects such as INTERACTIVE, SAFE-SPOT, and PREVENT deployed the advances in advanced driver-assistance systems (ADASs) and developed human-machine-interaction approaches for assisted driving. These European initiatives were accompanied by national research programs. In Germany, examples of these include MOTIV (1996–2000), INVENT (2001–2006) AKTIV (2006–2010), SIM^{TD} (2008–2013) and UR:BAN (2012–2016). A major achievement of these projects was intense cooperation between European OEMs, suppliers, and university researchers. Similarly, there has been a series of ASV (Advance Safety Vehicle) projects in Japan (ASV1 (1991–1995), ASV2 (1996–2000), ASV3 (2001–2005), ASV4 (2006–2010), ASV5 (2011–)), involving collaboration between the government and car manufacturers.

As part of the research on ADAS, there were a number of new measures of driving performance developed for car following (time headway, THW) [121, 122], lane keeping (time-to-line crossing, TLC) [123, 124], and braking maneuvers (time to collision, TTC) [125–127] over this decade and prior decades.

If several ADASs are installed in a vehicle, various warnings and other information will be given to the driver. In complex driving situations, multiple warning signals may occur simultaneously. In such cases, the driver may become confused and be unable to respond to the warnings or may not react appropriately. Therefore, warning signals should be integrated (ISO TR 12204) [109, 128].

An important human factors element of ADAS as assistive technology is the relationship between the driver and the system and especially the human-machine interface. In many cases, the driver receives feedback on the system state via the speedometer-tachometer cluster, center console displays, or a HUD, supplemented by force feedback from the steering wheel and pedals. If the driver does not comprehend the system's behavior as it actually is, "automation surprise" occurs when the system behaves unexpectedly. Therefore, interaction concepts for these systems have to take into account phenomena such as "over trust" and "over reliance" on the system to avoid serious problems [129]. Currently, numerous ADASs are available as mature products to support longitudinal and lateral vehicle control. Over time, the control authority of these systems has increased, and more complex, cooperative systems have been investigated [130–132]. How to integrate several ADASs and driver information systems has also been the topic of research [133, 134].

By definition ADASs are intended to assist drivers, so these systems must be designed to be compatible with driving behavior. An ADAS that does not consider driver ergonomic requirements may increase the risk of a crash, even though its aim is to enhance safety. Human factors research is necessary to understand how drivers behave with or without the systems in an actual road environment, not in a laboratory experiment. The research methods described in the next section are necessary for such research.

4.3. Naturalistic-Driving Studies and Driving Simulator Studies. One of the research developments of the 1990s has been the completion of several naturalistic-driving studies as knowing what normally happens on real roads is necessary when developing ADASs. If driving situations are known to be dangerous, then the type of ADAS that should be developed for safety assistance is readily determined. Also, quantitative analysis of driving behavior on actual roads is beneficial for developing vehicle-safety technologies, as well as for developing future driver-assistance systems.

Traditional human factors studies involving controlled experiments are relatively low cost. On the other hand, one cannot conduct a naturalistic-driving experiment of any size for less than \$10,000,000, and many cost much more. For that price, one could conduct 20–100 driving simulator experiments, depending upon their complexity. Until the 1990s, there was not sufficient interest in the topics that naturalistic-driving studies address to find funding for them.

Second, naturalistic-driving studies require compact data-collection hardware, low power, a large amount of data-storage capability, and sophisticated wireless communication, so that highly reliable and readily accessed data can be collected. Before the 1990s, that technology did not exist.

In the USA, the National Highway Traffic Safety Administration (NHTSA) conducted the 100-Car Naturalistic-Driving Study in 2001. They collected data on vehicle behavior, road-traffic conditions, and driver behavior in accidents and near-accident incidents, using vehicle-acceleration data as the trigger for recording. This study demonstrated that various distracting situations lead to traffic accidents in the real world [135]. Other relevant studies include the Advanced Collision Avoidance System (ACAS) [136], RDCW [137], and IVBSS [138] projects. Easily installed driving recorders for general-use passenger vehicles became commercially available in Japan in 2003. Detailed causal analysis of accidents and near accidents became possible with this device. JSAE has examined data gathered by driving recorders installed in taxis [139]. They conducted statistical analyses to classify the causes of accidents and also identified specific situations in which drivers committed behavioral errors.

The New Energy and Industrial Technology Development Organization (NEDO) of Japan conducted a three-year project beginning in 2001 to collect driving-behavior data under normal conditions, with no accidents, using instrumented vehicles in real road environments, and compiled the results in a database. This driving-behavior database has been publicly available since 2004 and has been used by universities, research institutions, and industry in research and development activities [140].

In Europe, the EURO-FOT (field operational test) study and the Promoting real Life Observations for Gaining Understanding of road user behaviour in Europe (PROLOGUE) project gathered remarkable naturalistic-driving datasets. EURO-FOT focused on the usage patterns considering ADAS and driver information system applications. An important output from EURO-FOT was the so-called FESTA handbook [141], which provides good practice recommendations for conducting naturalistic-driving studies.

For ADASs that assist steering and pedal control, a control algorithm should be developed to match the driver expectations of the system's behavior. If the control algorithm of ADAS is different from what the driver expects, the driver may feel uneasy and may not use the system. Traditional research methods, designed to repeat a controlled set of conditions so that they can be examined in a cost efficient manner, are an imperfect representation of the real world. For ADAS design, addition information was obtained from naturalistic-driving-behavior studies [142, 143]. To analyze the large data sets from naturalistic-driving studies, specialized statistical-modeling techniques are used [144, 145]. However, prior to applying these methods, the situations and conditions in which the targeted driving behavior occurs must be identified to create the subset of the data desired for analysis.

Improvements in computer-graphics technology and computing performance enabled detailed representation of road structures and traffic-participant behavior. As a result, simulators could be used as tools in driver-behavior research. Using a driving simulator, experiments can be conducted repeatedly, controlling such traffic situations as the positions of other vehicles relative to the subject vehicle. Experiments using a driving simulator are time efficient and do not expose subjects to the risk of real injury in a crash. Taking advantage of these capabilities, researchers are able to analyze the effectiveness of systems being developed and can anticipate potential problems by analyzing drivers' responses to the prototypes [146].

Until the 1990s, driving simulators were only found in a limited number of laboratories, primarily because of their cost. In part this was because rendering of scenes required high-performance graphic processors, and prior to the 1990s systems with adequate performance were specialized and costly. Second, projectors that had adequate resolution and brightness were also quite costly. After the '90s, the simulator hardware components became much less expensive.

Simulators are useful tools for investigating driver behavior. Driving simulators range from those resembling PC games to full-scale driving simulators such as the National Advanced Driving Simulator (NADS) and Toyota driving simulators. Although driving simulators are now commonly used for automotive human factors research, the research must be conducted with a clear understanding of what each simulator is capable of reproducing and to what degree, and with sufficient assessment or validation of the appropriateness of use for the experiment's purpose [147]. New researchers often do not spend enough time to make sure the values of the dependent measures collected are reasonable for real vehicles. A high-quality research program will likely include a balance of simulator experiments and actual road experiments or naturalistic-driving data [148].

Naturalistic-driving studies and driving simulator studies have proven to be powerful tools for analyzing driving behavior, assessing effectiveness, and identifying problems not only of driver information but also of driver-assistance systems. In the past, automotive human factors research typically focused on the relationship between drivers and vehicles. Now, research has gone beyond the human factors laboratories and extended to human behavioral research in the real world.

4.4. Driver Communications External to the Vehicle—Network Service, Mobile Phones, and Internet Access While Driving.

The introduction of information-communication technology has been particularly important for driver information systems. The Vehicle Information and Communication System (VICS), which transmits real-time traffic conditions for specific driving regions through FM radio signals and radio/optical beacons, began operating in 1996 in Japan. In Europe, the Radio Data System (RDS), introduced in the 1980s, later became the Traffic Message Channel (RDS-TMC); it conveys traffic information and messages via the FM signal. OnStar, a network service for GM, was started in 1995 in the USA. This was followed by TeleAid in 1997 in Germany, Toyota's MONET in 1997, Nissan Carwings in 1998 in Japan and BMW Assist and Mercedes MBrace in the late '90s. When the driver accesses the remote operations center of one of these systems, the operator assists with the trip according to the driver's request. Analysis of verbal communication between the driver and the operator, such as phrases used, the timing of utterances and pauses, and the number of turns, will provide insights into designing interactive speech systems for driver information systems.

Mobile radio phones installed in vehicles were first developed in 1947 by AT&T, but the service area was very limited and the phone itself was bulky. The A-Netz mobile-phone network started in Germany in 1952. The first cellular network began operating in 1979 in Japan. In the mid-1990s, cellular phones spread rapidly based on the Global System for Mobile communications (GSM) standard, and, not surprisingly, people used the phone while driving. The use of cellular phones while driving soon became a public-safety concern, and using a handheld cellular phone while driving was forbidden in many European member states in the 1990s, in Switzerland in 1996, and in Japan in 1999 [102]. Use of cellular phones for conversation is also illegal in some states in the USA [http://www.ncsl.org/issues-research/transport/cellular-phone-use-and-texting-while-driving-laws.aspx] and in a number of countries [149]. Hands-free systems for vehicles have since been introduced and have been shown to be less distracting [150]. The nature and extent of the interference of phone conversations while driving continues to be an important research topic [135, 151–154]. Of increasing importance is the effect of using cell phones on situation awareness [155]. Nonetheless, people use phones while driving for many reasons: they may feel that they do not have too much to do, believe driving is wasted time, feel a need to be connected, are bored, or for many other reasons. Use of phones while driving is widespread [156].

Voice communication by phone is one of many ways for people to communicate and interact with each other and with information systems. However, if the in-vehicle system restricts the access to information strictly for safety purposes, drivers might not connect the device to the in-vehicle system, bypassing the restrictions imposed by the vehicle. How to support interaction with data in these devices that drivers need and want while driving without relying on a visual-manual interface needs further human factors research. Interestingly, relative to the amount of research on phone use in conversation, relatively little research has been done on interaction with the Internet and intelligent systems while driving [157].

Some thought should also be given to what drivers really want or need to know. Qualitative methods for recording and analyzing human behavior in daily life are being developed in the field of sociology [158–160]. Such methods include ethnography, which describes detailed human behavior, and action research, in which the researcher explores problems of a society while acting as a member of the targeted society (See also [161, 162]).

Communication with those outside the vehicle that is not relevant to the driving task can cause driver distraction. Compared with interactions with driver information systems or ADASs, communication through mobile phones and the Internet is independent of the driving itself. Incoming alerts for phone calls, e-mail, and Short Message Service are external system-initiated interactions that occur regardless of the driving situation. There is a basic potential of driver distraction. To avoid this, there is a big potential if communication devices (nomadic devices) are connected to an in-vehicle information system that can control interaction with the driver to support the driver in the management of his workload. Discussions of possible mechanisms and interfaces for managing information to reduce workload and enhance situation awareness of the drivers were reported in the ITU-T FG Distraction activity [163–165]. However, the nomadic device should first be connected to the in-vehicle system, but should not bother the user. Connectivity technologies such as Bluetooth are important enablers here. Human factors research must design the in-vehicle system to give the driver an incentive to connect the device. Targets of human factors research are not only reducing workload and improving ease of use, but also designing system to induce safe driving.

4.5. Vehicle Communications with Other Vehicles and the Infrastructure. At first thought, these communications would appear to have nothing to do with human factors, which would be incorrect. The purpose of these communications is ultimately to deliver information to the driver. A major part of the cost of building systems to warn of and avoid crashes is the sensing systems, the radar, LIDAR, video, and sonar technologies to provide 360 degree coverage to support the driver. These sensors provide information to determine where all the threats are to the vehicle. This requires identifying each target from the background, identifying the type of target it is, and then developing a prediction of its path, which is used to determine if the target will collide with the driven vehicle. For locations where crashes are frequent, embedding sensors into the infrastructure is cost effective. Infrastructure-based cooperative systems were developed in Automated Highway System (AHS) projects (1996–2007) and the Driving Safety Support System (DSSS) project in Japan [166]. DSSS became operational in 2011 as a pilot study [167]. The system detects vehicles that are hidden by road structures at intersections, merging zones, and curves and informs the driver using an in-vehicle display and by voice [168]. An alternative approach is being examined under the UMTRI-led Safety Pilot program [169] and in other connected-vehicle activities. In a connected-vehicle approach, every vehicle, every pedestrian, and some key fixed objects that are part of the road infrastructure continuously transmit radio signals that communicate what they are, where they are, and, if they are capable of moving, how fast and in what direction they are moving. This, when fully fielded, could simplify the collision detection problem and lead to a potentially significant reduction in crashes, if the response to potential collisions is automatic.

What remains unknown is how to get drivers to respond to hazards they cannot see and may not become an imminent threat for some time [170]. How drivers should be warned if some of the broad array of information is unavailable, and when vehicles should take over the primary driving task will be a focus of future human factors research.

4.6. Autonomous Vehicles—Removing the Driver from Control. Until recently, self-driving cars seemed like a futuristic concept. However, with DARPA's Grand Challenge program [171], Google's demonstrations (http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works), and other activities such as Stadtpilot in Germany [172], advances in autonomous vehicles are occurring quickly.

Questions of concern to human factors researchers include the following: When can automation do a better job of driving than a human being? How can drivers be kept informed of the driving situation? How does the hand-over (driver to vehicle, vehicle to driver) occur? How do drivers of nonautonomous vehicles negotiate with the behavior of autonomous vehicles?

5. What Can Be Learned from History?

In general, the introduction of the automobile and the related achievements in human factors can be called a success story, having served as a stimulus for other research domains.

(1) Over time, the human factors focus has shifted from relying on personal experience to relying on research data that eventually led to standards from SAE, ISO, and others. However, as vehicles evolve, there will continue to be a need to conduct research to develop new standards, and to support the design of vehicles. Relative to other fields of engineering, the use of models to predict human performance while driving (except for control theory and workspace layout) has been limited [173]. Research on computational models of the heterogeneous group of drivers as information processors in very different traffic situations is needed as well as a significant effort to build practical tools engineers can use [174, 175]. Given what has occurred in the past, an important step would be incorporating those models in SAE and/or ISO standards.

(2) Over time, the primary problems that human factors experts address have increasingly shifted from physical to cognitive, but the original problems never go away. Early human factors efforts concerned making sure that drivers

could operate controls while providing adequate force to steer and brake. Although power-assist systems have assured braking and steering can be accomplished, questions about the optional human-device transfer function remain, as well as where to place controls so they can be comfortably operated. There are still issues of field of view, seating comfort, and thermal comfort, especially in connection with electric vehicles. Designers still wrestle with these issues and continue to request better data, better models, and better tools.

(3) Over time, there has been a shift in what the driver does. Initially, the driver just steered the vehicle, sometimes assisted by the codriver. Now, the driver controls an array of information and communication systems being assisted by the vehicle. Driver distraction and overload are major concerns. Research on how to coordinate performing the primary driving task and communicate with those outside of the vehicle, or both people and vehicles, are needed. The need for driver assistance is continuously increasing, especially in urban settings.

(4) Over time, developments in the automotive industry related to human factors mirror technology developments in general with a shift from providing basic mobility to concerns about crash protection and fuel efficiency. The early developments were related to the physical structure of the vehicle, the province of the mechanical engineer. More recent developments are the province of electrical and computer engineers. The most recent efforts, such as the nomadic device forum of the AIDE project, have involved engineers who develop nomadic and mobile devices brought into the vehicle. The next phase of vehicle evolution may center on the motor vehicle as a social mechanism, thus involving urban planners, sociologists, anthropologists, and others. One example of this concerns how to support the use of social networks (and what should be supported) while driving.

(5) Over time, evaluation methods have changed. The original human factors work was based strictly on intuition. That was followed by decades of research involving single test vehicles in scripted on-road experiments along with the analysis of crash data, almost exclusively from the USA. In the last few decades, the use of driving simulators in combination with eye tracking, but also laboratory evaluation of interaction concepts, has become much more widespread. The major recent development in methods has been naturalistic-driving studies and field operational tests, providing extensive real-world driving data. What remains unknown is at what point these studies transition from independent evaluations to a continuing data collection effort analogous to crash evaluations. Also unknown is when some country other than the USA will make its crash data publically available on the web. Without such information, research and design solutions will invariably focus on American problems, which may not match the driving situation in other countries.

(6) Over time, the way in which designers and researchers interact has changed. Initially, that occurred though major, large conferences such as the SAE Annual Congress, the TRB Annual Meeting, and others. Increasingly, however, the preferred venues are smaller, more focused meetings concerning automotive human factors in general, or specific aspects of that topic such as Driving Assessment and AutoUI. In addition, an important degree of informal interaction occurs at standardization meetings of various types.

(7) The news media have been a significant factor in bringing issues of crash safety to light, at least in the USA. Fires, crashes in which children are killed, and rollovers invariably get the most attention. At least once every decade there are major questions raised about the safety of at least one vehicle—Chevrolet Corvair, Ford Pinto, GM CK pickup trucks, Jeep CJ-5, Audi 5000, Ford Crown Victoria, Ford Explorer, and so forth. As a result, auto sales plummet for these models, and the manufacturers respond. Not all of the problems receiving attention from them have been genuine. However, at least in the USA, laws have been passed, research funded, and organizations created because of these media investigations.

The role of the news media in the future is uncertain. The USA was traditionally dominated by three television networks—NBC, ABC, and CBS. However, in recent years there has been competition from other networks in the USA, and foreign networks will soon have a greater presence in the USA. The competition has reduced funding for investigative journalism, but in its place, Internet journalism has arisen.

(8) Until now, automotive research and design have been dominated by the USA, Europe, and Japan. However, with China being the largest market for motor vehicles, and a growing market in India, there is the potential for them to be leading contributors to the automotive human factors research and design in the future.

Thus, although many may view traditional motor vehicles as part of an outdated industry, in fact, the industry has continued to evolve, with continuing pressure to introduce new technology into vehicles to increase safety and comfort and to develop cleaner, more fuel-efficient vehicles. However, the challenge the motor vehicle industry faces that the consumer products industry does not face is the high level of reliability and durability required, a concern that dates back decades ago as described in the literature.

As one can tell from the references provided, there has been an abundant and almost overwhelming amount of research conducted on automotive human factors. Those wishing to delve more deeply into the field may wish to begin by considering other overviews of automotive human factors, such as [176–182]. As the field of automotive human factors continues to evolve, it is important for designers, engineers, researchers, and others working on this topic to continue to learn about it. Reading a few papers or taking a human factors class is not enough. To keep informed, one needs to continue reading about the field, attend conferences, and participate in professional activities.

Acknowledgment

Copyright (c) 1956 SAE International. Reprinted with permission from SAE 560061/SP-142A.

References

[1] F. W. Wells, *Occupant Protection and Automobile Safety in the U.S. Since 1900*, SAE International, Warrendale, Pa, USA, 2012.

[2] A. J. Yanik, "The first 100 years of transportation safety: part 1," in *The Automobile: A Century of Progress*, pp. 121–132, Society of Automotive Engineers, Warrendale, Pa, USA, 1997.

[3] "25 Years Progress of SAE," A leaflet of the 25th Anniversary Celebration of the Society of Automotive Engineers, Society of Automotive Engineers, Warrendale, Pa, USA, May 1930.

[4] W. Kamm, *Das Kraftfahrzeug*, Springer, Berlin, Germany, 1936.

[5] A. J. Yanik, "The first 100 years of transportation safety: part 2," in *The Automobile: A Century of Progress*, pp. 133–149, Society of Automotive Engineers, Warrendale, Pa, USA, 1997.

[6] N. Ach, "Psychologie und technik bei bekämpfung von autounfällen," *Industrielle Psychotechnik*, vol. 6, no. 3, pp. 87–105, 1929.

[7] T. W. Forbes, "The normal automobile driver as a traffic problem," *The Journal of General Psychology*, vol. 20, pp. 471–474, 1939.

[8] D. Meister, *The History of Human Factors and Ergonomics*, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 1999.

[9] H. W. Sinaiko, *Selected Papers on Human Factors in the Design and Use of Control Systems*, Dover, Mineola, NY, USA, 2000.

[10] A. Chapanis, W. R. Garner, and C. T. Morgan, *Applied Experimental Psychology*, John Wiley & Sons, New York, NY, USA, 1949.

[11] National Research Council, *Human Factors in Undersea Warfare*, National Research Council, Committee on Undersea Warfare, Panel on Psychology and Physiology, Washington, DC, USA, 1949.

[12] P. G. Ronco, "Human factors engineering, bibliographic series volume 1 1940–1959 literature," Technical Report AD 639806, Tufts University, Medford, Mass, USA, 1966.

[13] R. A. McFarland and H. W. Stoudt, "Human body size and passenger vehicle design," SAE Special Publication 142, Society of Automotive Engineers, Warrendale, Pa, USA, 1955.

[14] J. Kulowski, "Orthopedic aspects of automobile crash injuries and deaths," *Journal of the American Medical Association*, vol. 163, no. 4, pp. 230–233, 1957.

[15] E. R. Dye, "Kinematics of the human body under crash conditions," *Clinical Orthopaedics*, vol. 8, pp. 305–309, 1956.

[16] "Manikins for use in defining vehicle seating accommodation," SAE Recommended Practice J826, 1962.

[17] "Motor vehicle driver's eye range," SAE Recommended Practice J941, 1965.

[18] D. Hammond and R. Roe, "Driver head and eye positions," SAE Technical Paper 720200, Society of Automotive Engineers, Warrendale, Pa, USA, 1972.

[19] J. F. Meldrum, "Automobile driver eye position," SAE Technical Paper 650464, Society of Automotive Engineers, Warrendale, Pa, USA, 1972.

[20] "Passenger car rear vision," SAE Recommended Practice J834, 1967.

[21] R. Roe and P. Kyropoulos, "The application of anthropometry to automotive design," SAE Technical Paper 70053, Society of Automotive Engineers, Warrendale, Pa, USA, 1970.

[22] Occupant Crash Protection, "Federal Motor Vehicle Safety Standard (FMVSS) 208, 49 CFR 571.208," Standard 208, 1959.

[23] E. Simonson, C. Baker, N. Burns, C. Keiper, O. H. Schmitt, and S. Stackhouse, "Cardiovascular stress (electrocardiographic changes) produced by driving an automobile," *American Heart Journal*, vol. 75, no. 1, pp. 125–135, 1968.

[24] R. C. Jagacinski and J. M. Flach, *Control Theory For Humans*, Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2003.

[25] D. McRuer and D. H. Weir, "Theory of manual vehicular control," *Ergonomics*, vol. 12, no. 4, pp. 599–633, 1969.

[26] T. B. Sheridan, Ed., *Mathematical Models and Simulation of Automobile Driving*, Conference Proceedings, Massachusetts Institute of Technology, Cambridge, Mass, USA, 1967.

[27] N. Rashevsky, "Man-machine interaction in automobile driving," *Progress in Biocybernetics*, vol. 42, pp. 188–200, 1964.

[28] C. C. MacAdam, "Application of an optimal preview control for simulation of closed-loop automobile driving," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 11, no. 6, pp. 393–399, 1981.

[29] I. D. Brown and E. C. Poulton, "Measuring the spare "mental capacity" of car drivers by a subsidiary task," *Ergonomics*, vol. 4, no. 1, pp. 35–40, 1961.

[30] J. W. Senders, A. B. Kristofferson, W. H. Levison, C. W. Dietrich, and J. L. Ward, "The attentional demand of automobile driving," Highway Research Record 195, 1967.

[31] R. E. Beinke and J. K. Williams, "Driving simulator," in *Proceedings of Automotive Safety Seminar*, vol. 24, General Motors, Warren, Mich, USA, 1968.

[32] E. Kikuchi, T. Matsumoto, S. Inomata, M. Masaki, T. Yatabe, and T. Hirose, "Development and application of high speed automobile driving simulator," Technical Report of Mechanical Engineering Laboratory 89, 1976 (Japanese).

[33] J. Drosdol and F. Panik, "The Daimler-Benz driving simulator: a tool for vehicle development," SAE Technical Paper 850334, Society of Automotive Engineers, Warrendale, Pa, USA, 1985.

[34] R. Nader, *Unsafe at any Speed*, Grossman, New York, NY, USA, 1965.

[35] G. T. Schwartz, "The myth of the Ford Pinto case," *Rutgers Law Review*, vol. 43, pp. 1013–1068, 1991.

[36] M. Dowie, "Pinto madness," *Mother Jones*, September-October 1977.

[37] W. M. Hoffman, "Case study—the Ford Pinto," *Corporate Obligations and Responsibilities: Everything Old is New Again*, 222–229, 1966.

[38] A. Irving and K. S. Rutley, "Some driving aids and their assessments," in *Proceedings of the Symposium on Psychological Aspects of Driver Behavior*, Institute for Road Safety Research, 1971.

[39] R. E. Reilly, D. S. Kurke, and C. C. Buckenmaier, "Validation of the reduction of rear-end collisions by a high-mounted auxiliary stop lamp," Technical Report DOT HS 805 360, U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC, USA, 1980.

[40] P. L. Olson, "Evaluation of a new LED high-mounted stop lamp, in vehicle lighting trends," Special Publication SP-692, Society of Automotive Engineers, Warrendale, Pa, USA, 1987.

[41] K. Rumar, G. Helmers, and M. Thorell, "Obstacle visibility with European Halogen H4 and American sealed beam headlights," Tech. Rep. 133, University of Uppsala, Department of Psychology, Uppsala, Sweden, 1973.

[42] J. K. Foster, J. D. Kortge, and M. J. Wolanin, "Hybrid III-A biomechanically-based crash test dummy," SAE Technical Paper 770938, Society of Automotive Engineers, Warrendale, Pa, USA, 1977.

[43] J. Versace, "A review of the severity index," SAE Technical Paper 710881, Society of Automotive Engineers, Warrendale, Pa, USA, 1971.

[44] D. C. Hammond and R. W. Roe, "SAE controls reach study," SAE Technical Paper 720199, Society of Automotive Engineers, Warrendale, Pa, USA, 1972.

[45] "Driver hand control reach," SAE Recommended Practice J287, 1976.

[46] "Direction-of-motion stereotypes for automotive hand controls," SAE Recommended Practice J1139, 1977.

[47] "Symbols for motor vehicle controls," SAE Standard J1048, 1974.

[48] M. J. Griffin, *Handbook of Human Vibration*, Elsevier, London, UK, 1996.

[49] "Mechanical vibration and shock—guide for the evaluation of human exposure to whole-body vibration," ISO 2631, 1974.

[50] R. R. Mourant and T. H. Rockwell, "Mapping eye-movement patterns to the visual scene in driving: an exploratory study," *Human Factors*, vol. 12, no. 1, pp. 81–87, 1970.

[51] B. Richter, "Driving simulator studies: the influence of vehicle parameters on safety in critical situations," SAE Technical Paper 741105, Society of Automotive Engineers, Warrendale, Pa, USA, 1974.

[52] R. C. McLane and W. W. Wierwille, "The influence of motion and audio cues on driver performance in an automobile simulator," *Human Factors*, vol. 17, no. 5, pp. 488–501, 1975.

[53] R. G. Snyder, T. L. McDole, W. M. Ladd, and D. J. Minahan, "On-road crash experience of utility vehicles," Tech. Rep. UM-HSRI-80-14, Highway Safety Research Institute, Ann Arbor, Mich, USA, 1980.

[54] R. G. Snyder, T. L. McDole, W. M. Ladd, and D. J. Minahan, "An overview of the on-road crash experience of utility vehicles (highlights of the technical report)," Tech. Rep. UM-HSRI-80-15, Highway Safety Research Institute, Ann Arbor, Mich, USA, 1980.

[55] S. Franklin and M. Stepanek, *Trouble in Jeep Country, AMC Claima It's Safe*, Detroit Free Press, 1983.

[56] P. Niedermeyer, "The Best of TTAC: The Audi 5000 Intended Unintended Acceleration Debacle," http://www.thetruthaboutcars.com/2010/03/the-best-of-ttac-the-audi-5000-intended-unintended-acceleration-debacle/.

[57] R. Walter, G. Carr, H. Weinstock, D. Sussman, and J. Pollard, "Study of mechanical and driver-related systems of the Audi 5000 capable of producing uncontrolled sudden acceleration incidents," Tech. Rep. DOT-TSC-NHTSA-88-4, U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC, USA, 1988.

[58] National Aeronautics and Space Administration, Technical Support to the National Highway Traffic Safety Administration (NHTSA) on the Reported Toyota Motor Corporation (TMC) Unintended Acceleration (UA) Investigation (NESC Assessment TI-10-00618), National Aeronautics and Space Administration, NASA Safety and Engineering Center, January 2011.

[59] J. D. Brown, "The opportunities of ergonomics," in *Human Factors in Transport Research. Vehicle Factors: Transport Systems, Workspace, Information and Safety*, D. J. Oborne and J. A. Lewis, Eds., vol. 1, Academic Press, New York, NY, USA, 1980.

[60] H. Bubb, F. Engstler, F. Fritzsche et al., "The development of RAMSIS in past and future as an example for the cooperation between industry and university," *International Journal of Human Factors Modelling and Simulation*, vol. 1, no. 1, pp. 140–157, 2006.

[61] P. Blanchonette, "Jack human modelling tool: a review," Tech. Rep. DSTO-TR-2364, Defense Science and Technology Organization Victoria (Australia) Air Operations Division, Fishermans Bend, Victoria, Australia, 2010, document ADA 518132.

[62] K. S. Rutley, "Control of drivers' speed by means other than enforcement," *Ergonomics*, vol. 18, no. 1, pp. 89–100, 1975.

[63] K. Bengler, H. Bubb, I. Totzke, J. Schumann, and F. Flemisch, "Automotive," in *Information Ergonomics—A Theoretical Approach and Practical Experience in Transportation*, P. Sandle and M. Stein, Eds., Springer, Heidelberg, Germany, 2012.

[64] K. Parsons, *Human Thermal Environments*, Taylor & Francis, London, UK, 2nd edition, 2003.

[65] D. P. Gatley, "Psychrometric chart celebrate 100th anniversary," *ASHRAE Journal*, vol. 46, no. 11, pp. 16–20, 2004.

[66] D. P. Wyon, C. Tennstedt, I. Lundgren, and S. Larsson, "A new method for the detailed assessment of human heat balance in vehicles, Volvo's thermal manikin, VOLTMAN," SAE Technical Paper 850042, Society of Automotive Engineers, Warrendale, Pa, USA, 1985.

[67] C. W. Erwin, J. W. Hartwell, M. R. Volow, and G. S. Alberti, "Electrodermal change as a predictor of sleep," in *Studies of Drowsiness (Final Report)*, C. W. Erwin, Ed., The National Driving Center, Durham, North Carolina, 1976.

[68] T. A. Dingus, L. H. Hardee, and W. W. Wierwille, "Detection of drowsy and intoxicated drivers based on highway driving performance measures," IEOR Department Report #8402, Virginia Tech, Department of Industrial Engineering and Operations Research, Blacksburg, Va, USA, 1985.

[69] W. W. Wierwille, "Research on vehicle-based driver status/performance monitoring, development, validation, and refinement of algorithms for detection of driver drowsiness," Technical Report DOT HS 808 247, U.S. Department of Transportation, Washington, DC, USA, 1994.

[70] D. A. Spyker, S. P. Stackhouse, S. Khalalfalla, and R. C. McLane, "Development of techniques for measuring pilot workload," Contractor report NASA CR 1888, NASA, Washington, DC, USA, 1971.

[71] T. Miura, "Coping with situational demands: a study of eye movements and peripheral vision performance," in *Vision in Vehicles 1*, A. G. Gale, M. H. Freeman, C. M. Haslegrave, P. Smith, and S. P. Taylor, Eds., pp. 205–216, North-Holland, Amsterdam, The Netherlands, 1986.

[72] S. Nordomack, H. Jansson, M. Lidstrom, and G. Palmkvist, "A moving base driving simulator with wide angle visual system," VTI Technical Report 106A, Swedish Road and Traffic Research Institute, Linkoping, Sweden, 1986.

[73] S. Hahn and W. Kaeding, "The Daimler-Benz driving simulator-presentation of selected experiments," SAE Technical Paper 880058, Society of Automotive Engineers, Warrendale, Pa, USA, 1988.

[74] E. Blana, "A survey of driving research simulators around the world," ITS Working Paper 481, University of Leeds, Institute for Transport Studies, Leeds, UK, 1996.

[75] C. Y. D. Yang, J. D. Fricker, and T. Kuczek, "Designing advanced traveler information systems from a driver's perspective: Results of a driving simulation study," *Transportation Research Record*, no. 1621, pp. 20–26, 1998.

[76] W. Janssen and R. van der Horst, "Presenting descriptive information in variable message signing," *Transportation Research Record*, no. 1403, pp. 83–87, 1993.

[77] M. P. Reed and P. A. Green, "Comparison of driving performance on-road and in a low-cost simulator using a concurrent telephone dialling task," *Ergonomics*, vol. 42, no. 8, pp. 1015–1037, 1999.

[78] A. Steinfeld and P. Green, "Driver responses to navigation information on full-windshield, head-up displays," *International Journal of Vehicle Design*, vol. 19, no. 2, pp. 135–149, 1998.

[79] J. R. Bloomfield, J. R. Buck, J. M. Christensen, and A. Yenamandra, "Human factors aspects of the transfer of control from the driver to the automated highway system," Tech. Rep. FHWA-RD-94-173, U.S. Department of Transportation, Federal Highway Administration, Washington, DC, USA, 1994.

[80] T. Suetomi and K. Kido, "Driver behavior under a collision warning system: a driving simulator study," SAE Technical Paper 970279, Society of Automotive Engineers, Warrendale, Pa, USA, 1997.

[81] J. Godthelp and J. Schumann, "The use of an intelligent accelerator as an element of a driver support system," in *Proceedings of the 24th ISATA International Symposium on Automotive Technology and Automation*, 1991.

[82] W. W. Wierwille, M. C. Hulse, T. J. Fischer, and T. A. Dingus, "Strategic use of visual resources by the driver while navigating with an in-car navigation display system," SAE Technical Paper 885180, Society of Automotive Engineers, Warrendale, Pa, USA, 1988.

[83] H. K. Zwahlen, in *Information Processing, Driver Performance Data Book, Technical Report DOT HS 807 121*, R. L. Henderson, Ed., U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC, USA, 1987.

[84] F. J. Mammano and R. Favout, "An electronic route-guidance system for highway vehicles," *IEEE Transactions on Vehicular Technology*, vol. 19, no. 1, pp. 143–152, 1999.

[85] H. Ito, "Integrated development of automotive navigation and route guidance system—product development for realization of dreams and standardization," *Synthesiology (English Edition)*, vol. 4, no. 3, pp. 162–171, 2012.

[86] E. J. Blum, R. Haller, and G. Nirschl, "Driver-copilot interaction: modelling aspects and techniques," in *Proceedings of the 2nd Prometheus Workshop*, FHG-IITB, Stockholm, Sweden, 1989.

[87] J. A. Michon, *Generic Intelligent Driver Support*, Taylor & Francis, London, UK, 1993.

[88] A. M. Parkes and S. Franzen, *Driving Future Vehicles*, Taylor & Francis, London, UK, 1993.

[89] H. Ikeda, Y. Kobayashi, and K. Hirano, "How car navigation systems have been put into practical use: development management and commercialization process," *Synthesiology (English Edition)*, vol. 3, no. 4, pp. 280–289, 2011.

[90] P. Frame, "How GM trumped Dateline, maker used vast resources to dismantle pickup story," *Automotive News*, pp. 1, 1993.

[91] P. Green, W. Levison, G. Paelke, and C. Serafin, "Preliminary human factors guidelines for driver information systems," Tech. Rep. FHWA-RD-94-087, U.S. Department of Transportation, Federal Highway Administration, McLean, Va, USA, 1995.

[92] A. Stevens, A. C. Board, P. Allen, and A. Quimby, "A safety checklist of the assessment of in-vehicle information systems: scoring proforma," Project Report PA3536-A/99, Transport Research Laboratory, Crowthorne, UK, 1999.

[93] T. Ross, K. Midtland, M. Fuchs et al., *HARDIE Design Guidelines Handbook: Human Factors Guidelines for Information Presentation by ATT Systems*, Commission of the European Communities, Brussels, Luxembourg, 1996.

[94] P. Green, "Driver interface safety and usability standards: an overview," in *Driver Distraction Theory, Effects, and Mitigation*, M. Regan, J. Lee, and K. Young, Eds., pp. 445–464, CRC Press, Boca Raton, Fla, USA, 2008.

[95] "Wirtschaftsforum VerkehrstelematikVereinbarung zu Leitlinien für die Gestaltung und Installation von Informations- und Kommunikationssystemen in Kraftfahrzeugen," (English translation: Steering Group on the Economic Forum on Telematics in Transport, Agreement on guidelines for the design and installation of information and communication systems in motor vehicles), Bonn, Germany, 1996.

[96] O. M. J. Carsten and L. Nilsson, "Safety assessment of driver assistance systems," *European Journal of Transport and Infrastructure Research*, vol. 1, no. 3, pp. 225–243, 2001.

[97] C. Heinrich, "Automotive HMI International Standards," in *Proceedings 4th International Conference on Applied Human Factors and Ergonomics (AHFE '12)*, 2012.

[98] W. W. Wierwille, J. F. Antin, T. A. Dingus, and M. C. Hulse, "Visual attentional demand of a in-car navigational display system," in *Vision in Vehicles*, A. G. Gale, M. H. Freeman, C. M. Haslegrave, P. Smith, and S. P. Taylor, Eds., vol. 2, pp. 307–316, Elsevier, Amsterdam, The Netherlands, 1988.

[99] H. T. Zwahlen, C. C. Adams Jr., and D. P. DeBald, "Safety aspects of CRT touch panel controls in automobiles," in *Vision in Vehicles*, A. G. Gale, M. H. Freeman, C. M. Haslegrave, P. Smith, and S. P. Taylor, Eds., vol. 2, pp. 335–344, Elsevier, Amsterdam, The Netherlands, 1988.

[100] A. R. A. van der Horst, "Occlusion as a measure for visual workload: an overview of TNO occlusion research in car driving," *Applied Ergonomics*, vol. 35, no. 3, pp. 189–196, 2004.

[101] J. F. Krems, A. Keinath, M. Baumann, C. Gelau, and K. Bengler, "Evaluating visual display designs in vehicles: advantages and disadvantages of the occlusion technique," in *Advances in Network Enterprises, Virtual Organizations, Balanced Automation, and Systems Integration*, L. M. Camarinha-Matos, H. Afsarmanesh, and H. H. Erbe, Eds., pp. 361–368, Kluwer Academic, Norwell, Mass, USA, 2000.

[102] M. Akamatsu, "Japanese approaches to principles, codes, guidelines and checklists for in-vehicle HMI," in *Driver Distraction Theory, Effects, and Mitigation*, M. Regan, J. Lee, and K. Young, Eds., pp. 425–444, CRC Press, Boca Raton, Fla, USA, 2008.

[103] "Ergonomic aspects of transport information and control systems—occlusion method to assess visual distraction," ISO 16673, 2007.

[104] P. Green, "The 15-second rule for driver information systems," in *Proceedings of the ITS America 9th Annual Meeting*, ITS America, Washington, DC, USA, 1999.

[105] P. Green, "Estimating compliance with the 15-second rule for driver-interface bility and safety," in *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*, Santa Monica, Calif, USA, 1999.

[106] A. Baron and P. Green, "Safety and usability of speech interfaces for in-vehicle tasks while driving: a brief literature review," Technical Report UMTRI 2006-5, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 2006.

[107] V. E. Lo and P. Green, "Development and evaluation of automotive speech interfaces: useful information from the human factors and related literature," *International Journal of Vehicular Technology*, vol. 2013, Article ID 924170, 13 pages, 2013.

[108] "Ergonomic aspects of transport information and control systems—dialogue management principles," ISO 15005, 2002.

[109] "Ergonomic aspects of transport information and control systems—criteria for determining priority of messages," ISO TR16951, 2004.

[110] "Ergonomic aspects of transport information and control systems—visual presentation of information," ISO 15008, 2009.

[111] "Ergonomic aspects of transport information and control systems—suitability of TICS while driving," ISO 17287, 2003.

[112] "Ergonomic aspects of transport information and control systems—auditory information presentation," ISO 15006, 2004.

[113] "Ergonomic aspects of transport information and control systems—simulated lane change test to assess in-vehicle secondary task demand," ISO 26022, 2010.

[114] S. Mattes, "The lane-change-task as a tool for driver distraction evaluation," in *Quality of Work and Products in Enterprises of the Future*, H. Strasser, K. Kluth, H. Rausch, and H. Bubb, Eds., Erognomia, 2003.

[115] P. C. Burns, K. Bengler, and D. H. Weir, "Driver metrics and an overview of user needs and uses," in *Performance Metrics for Assessing Driver Distraction: The Quest for Improved Road Safety*, G. L. Rupp, Ed., chapter 1, pp. 24–30, SAE International, Warrendale, Pa, USA, 2010.

[116] A. Kumar, *Deadly Combination: Ford, Firestone and Florida*, Saint Petersburg Times, Saint Petersburg, Fla, USA, 2001.

[117] A. Kumar, *Attention Shi.s from Fires Tone to Ford Explorer*, Saint Petersburg Times, Saint Petersburg, Fla, USA, 2001.

[118] K. Naab and G. Reichart, "Driver assistance system for lateral and longitudinal vehicle guidance—heading control and active cruise support," in *Proceedings of International Symposium on Advanced Vehicle Control (AVEC '94)*, pp. S449–S454, Tsukuba, Japan.

[119] W. Prestl, T. Sauer, J. Steinle, and O. Tschernoster, "The BMW active cruise control ACC," SAE Technical Paper 2000-01-0344, Society of Automotive Engineers, Warrendale, Pa, USA.

[120] D. M. Hoedemaeker, *Driving with intelligent vehicles. Driving behaviour with adaptive cruise control and the acceptance by individual drivers [Ph.D. thesis]*, Delft University Press, Delft, The Netherlands, 1999.

[121] M. P. Heyes and R. Ashworth, "Further research on car-following models," *Transportation Research*, vol. 6, no. 3, pp. 287–291, 1972.

[122] P. Wasielewski, "Car following headways on freeways interpreted by the semi-Poisson headway distribution model," *Transportation Science*, vol. 13, no. 1, pp. 36–55, 1979.

[123] H. Godthelp, P. Milgram, and G. J. Blaauw, "The development of a time-related measure to describe driving strategy," *Human Factors*, vol. 26, no. 3, pp. 257–268, 1984.

[124] W. van Winsum, K. A. Brookhuis, and D. de Waard, "A comparison of different ways to approximate time-to-line crossing (TLC) during car driving," *Accident Analysis and Prevention*, vol. 32, no. 1, pp. 47–56, 2000.

[125] J. C. Hayward, "Near miss determination through use of a scale of danger," *Highway Research Record*, vol. 384, pp. 24–34, 1972.

[126] W. van Winsum and A. Heino, "Choice of time-headway in car-following and the role of time-to-collision information in braking," *Ergonomics*, vol. 39, no. 4, pp. 579–592, 1996.

[127] K. Vogel, "A comparison of headway and time to collision as safety indicators," *Accident Analysis and Prevention*, vol. 35, no. 3, pp. 427–433, 2003.

[128] "Ergonomic aspects of transport information and control systems—introduction to integrating safety critical and time critical warning signals," ISO TR 12204, 2012.

[129] I. Totzke, S. Jessberger, D. Mühlbacher, and H. P. Krüger, "Semi-autonomous advanced parking assists: do they really have to be learned if steering is automated?" in *Proceedings of European Conference on Human Centered Design for Intelligent Transport Systems*, pp. 123–132, Berlin, Germany, 2010.

[130] M. Kienle, D. Damböck, J. Kelsch, F. Flemisch, and K. Bengler, "Towards an H-Mode for highly automated vehicles: driving with side sticks," in *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive UI '09)*, pp. 19–23, September 2009.

[131] K. Bengler, M. Zimmermann, D. Bortot, M. Kienle, and D. Dambock, "Interaction principles for cooperative human-machine systems," *Information Technology*, vol. 54, no. 4, pp. 157–164, 2012.

[132] T. Inagaki and M. Itoh, "Human's overtrust in and overreliance on Advanced Driver Assistance Systems: a theoretical framework," *International Journal of Vehicular Technology*, vol. 2013, Article ID 951762, 8 pages, 2013.

[133] D. Popiv, C. Rommerskirchen, M. Rakic, M. Duschl, and K. Bengler, "Effects of assistance of anticipatory driving on driver's behaviour during deceleration situations," in *Proceedings of the 2nd European Conference on Human Centred Design of Intelligent Transport Systems (HUMANIST '10)*, Berlin, Germany, April 2010.

[134] D. Popiv, M. Rakic, F. Laquai, M. Duschl, and K. Bengler, "Reduction of fuel consumption by early anticipation and assistance of deceleration phases," in *Proceedings of the World Automotive Congress of International Federation of Automotive Engineering Societies (FISITA '10)*, Budapest, Hungary, June 2010.

[135] S. G. Klauer, T. A. Dingus, V. L. Neale, J. Sudweeks, and D. Ramsey, "The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car naturalistic driving study data," Technical Report DOT, HS 810 594, U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC, USA, 2006.

[136] R. Ervin, J. Sayer, D. LeBlanc et al., "Automotive collision avoidance system field operational test report: methodology and results," Technical Report HS 809 900, US Department of Transportation, National Highway Traffic Safety Administration, Washington, DC, USA, 2005.

[137] J. Sayer, C. Winkler, R. Ervin et al., "Road departure crash warning system field operational test: methodology and results. Volume 1: technical report," Tech. Rep. UMTRI-2006-9-1, U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC, USA, 2006.

[138] J. Sayer, D. LeBlanc, S. Bogard et al., "Integrated vehicle-based safety systems field operational test final program report," Technical Report HS 811 482, U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC, USA, 2011.

[139] M. Nagai, "Enhancing safety and security by incident analysis using drive recorders," *Review of Automotive Engineering*, vol. 27, no. 1, pp. 9–15, 2006.

[140] T. Sato and M. Akamatsu, "Influence of traffic conditions on driver behavior before making a right turn at an intersection: analysis of driver behavior based on measured data on an actual road," *Transportation Research F*, vol. 10, no. 5, pp. 397–413, 2007.

[141] "Field opErational teSt supporT Action (FESTA)," in *FESTA Handbook*, European Commission, Brussels, Belgium, 2013, http://www.its.leeds.ac.uk/festa/downloads/FESTA%20Handbook%20v2.pdf.

[142] T. Sato, M. Akamatsu, A. Takahashi et al., "Analysis of driver behaviour when overtaking with adaptive cruise control,"

Review of Automotive Engineering, vol. 26, no. 4, pp. 481–488, 2005.

[143] S. B. McLaughlin, J. M. Hankey, and T. A. Dingus, "A method for evaluating collision avoidance systems using naturalistic driving data," *Accident Analysis and Prevention*, vol. 40, no. 1, pp. 8–16, 2008.

[144] M. Akamatsu, Y. Sakaguchi, and M. Okuwa, "Modeling of driving behavior when approaching intersection based on measured behavioral data on an actual road," in *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*, pp. 1895–1899, 2003.

[145] T. Sato and M. Akamatsu, "Modeling and prediction of driver preparations for making a right turn based on vehicle velocity and traffic conditions while approaching an intersection," *Transportation Research F*, vol. 11, no. 4, pp. 242–258, 2008.

[146] J. D. Lee, D. V. McGehee, T. L. Brown, and M. L. Reyes, "Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator," *Human Factors*, vol. 44, no. 2, pp. 314–334, 2002.

[147] D. L. Fisher, M. Rizzo, J. Caird, and J. D. Lee, *Handbook of Driving Simulation*, CRC Press, Boca Raton, Fla, USA, 2011.

[148] T. Sato, M. Akamatsu, T. Shibata, S. Matsumoto, N. Hatakeyama, and K. Hayama, "Predicting driver behavior using field experiment data and driving simulator experiment data: Assessing impact of elimination of stop regulation at railway crossings," *International Journal of Vehicular Technology*, vol. 2013, Article ID 912860, 9 pages, 2013.

[149] K. L. Young and M. A. Regan, "Driver distraction exposure research: a summary of findings," in *Driver Distraction Theory, Effects, and Mitigation*, M. A. Regan, J. D. Lee, and K. L. Young, Eds., pp. 327–328, 2008.

[150] K. A. Brookhuis, G. de Vries, and D. de Waard, "The effects of mobile telephoning on driving performance," *Accident Analysis and Prevention*, vol. 23, no. 4, pp. 309–316, 1991.

[151] H. Alm and L. Nilsson, "Changes in driver behaviour as a function of handsfree mobile phones—a simulator study," *Accident Analysis and Prevention*, vol. 26, no. 4, pp. 441–451, 1994.

[152] W. J. Horrey and C. D. Wickens, "Examining the impact of cell phone conversations on driving using meta-analytic techniques," *Human Factors*, vol. 48, no. 1, pp. 196–205, 2006.

[153] D. Lamble, T. Kauranen, M. Laakso, and H. Summala, "Cognitive load and detection thresholds in car following situations: safety implications for using mobile (cellular) telephones while driving," *Accident Analysis and Prevention*, vol. 31, no. 6, pp. 617–623, 1999.

[154] D. L. Strayer and W. A. Johnston, "Driven to distraction: dual-task studies of simulated driving and conversing on a cellular telephone," *Psychological Science*, vol. 12, no. 6, pp. 462–466, 2001.

[155] B. Metz, N. Schömig, H. P. Krüger, and K. Bengler, "Situation awareness in driving with in-vehicle information systems," in *Performance Metrics for Assessing Driver Distraction: The Quest For Improved Road Safety*, G. L. Rupp, Ed., chapter 12, SAE International, Warrendale, Pa, USA, 2010.

[156] F. A. Drews and D. L. Strayer, "Cellular phone and driver distraction," in *Driver Distraction Theory, Effects, and Mitigation*, M. Regan, J. Lee, and K. Young, Eds., pp. 169–190, CRC Press, Boca Raton, Fla, USA, 2008.

[157] M. Vollrath, T. Meilinger, and H. P. Krüger, "How the presence of passengers influences the risk of a collision with another vehicle," *Accident Analysis and Prevention*, vol. 34, no. 5, pp. 649–654, 2002.

[158] J. R. Davis and C. M. Schmandt, "The back seat driver: real time spoken driving instructions," in *Proceedings of the IEEE Vehicle Navigation and Information Systems Conference (VNIS '89)*, pp. 146–150, September 1989.

[159] M. Akamatsu and M. Kitajima, "Designing products and services based on understanding human cognitive behavior—development of cognitive chrono-ethnography for synthesiological research," *Synthesiology (English Edition)*, vol. 4, no. 3, pp. 144–155, 2012.

[160] B. Brown, E. Laurier, H. Lorimer et al., "Driving and "passengering": Notes on the ordinary organization of car travel," *Mobilities*, vol. 3, no. 1, pp. 1–23, 2008.

[161] K. Bengler, J. F. Coughlin, B. Reimer, and B. Niedermaier, "A new method to investigate cognitive structures of user's on automotive functionalities," in *Proceedings of the 3rd International Conference on Applied Human Factors and Ergonomics (AHFE '10)*, Miami, Fla, USA, July 2010.

[162] B. Brown and E. Laurier, "The normal, natural troubles of driving with GPS," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, pp. 1621–1630, New York, NY, USA, 2012.

[163] ITU-T FG Distraction P.UIA report, FG Distraction report on User Interface requirements for Automotive applications (P.UIA).

[164] ITU-T FG Distraction G.V2A report, FG Distraction report on communications interface between external applications and a Vehicle Gateway Platform (G.V2A).

[165] ITU-T FG Distraction G.SAM report, FG Distraction report on Situational Awareness Management (G.SAM).

[166] H. Hatakenaka, H. Kanoshima, T. Aya, S. Nishi, H. Mizutani, and K. Nagano, "Development and verification of effectiveness of an AHS," in *Proceedings of ITS World Congress*, New York, NY, USA, 2008.

[167] M. Hatakeyama and S. Nakayama, "Progress toward the practical use of vehicle infrastructure cooperation system," in *Proceedings of ITS World Congress*, Stochholm, Sweden, 2009.

[168] K. Daimon, H. Makino, H. Mizutani, and Y. Munehiro, "Study on safety assist information of Advanced Cruise-Assist Highway Systems (AHS) using VICS in blind curve section of urban expressway," *Journal of Mechanical Systems For Transportation and Logistics*, vol. 1, no. 2, pp. 192–202, 2006.

[169] U. S. Department of Transportation, "Safety Pilot Program Overview," http://www.its.dot.gov/safety_pilot/index.htm.

[170] D. Popiv, C. Rommerskirchen, M. Rakic, M. Duschl, and K. Bengler, "Effects of assistance of anticipatory driving on driver's behaviour during deceleration situations," in *Proceedings of the 2nd European Conference on Human Centered Design of Intelligent Transport Systems (HUMANIST '10)*, Berlin, Germany, April 2010.

[171] S. Thrun, M. Monemerlo, H. Dahlkamp et al., "Stanley the robot that won the DARPA grand challenge," in *DARPA Grand Challenge: The Great Robot Race*, M. Buehler, K. Iagnemma, and S. Singh, Eds., vol. 36 of *Springer Tracts in Advanced Robotics*, pp. 1–43, 2007.

[172] T. Nothdurft, P. Hecker, S. Ohl et al., "Stadtpilot: first fully autonomous test drives in urban traffic," in *Proceedings of the 14th International IEEE Annual Conference on Intelligent Transportation Systems*, Washington, DC, USA, 2011.

[173] T. A. Ranney, “Models of driving behavior: a review of their evolution,” *Accident Analysis and Prevention*, vol. 26, no. 6, pp. 733–750, 1994.

[174] D. D. Salvucci, “Modeling driver behavior in a cognitive architecture,” *Human Factors*, vol. 48, no. 2, pp. 362–380, 2006.

[175] C. P. Cacciabue, Ed., *Modelling Driver Behaviour in Automotive Environment: Critical Issues in Driver Interactions with Intelligent Transport Systems*, Springer, London, UK, 2007.

[176] L. Evans, *Traffic Safety and the Driver*, Van Nostrand Reinhold, New York, NY, USA, 1991.

[177] B. Peacock and W. Karwowski, Eds., *Automotive Ergonomics*, Taylor & Francis, London, UK, 1993.

[178] M. Sivak, M. J. Flannagan, and B. Schoettle, “Driver assessment and training in the 1980s and 1990s: an analysis of the most-cited publications,” in *Proceedings of the Driving Assessment Conference*, 2001.

[179] J. D. Lee, “Fifty years of driving safety research,” *Human Factors*, vol. 50, no. 3, pp. 521–528, 2008.

[180] H. H. Braess and U. Seiffert, Eds., *Handbuch Kraftfahrzeugtechnik. 6. Auflage*, Vieweg and Teubner, Wiesbaden, Germany, 2011.

[181] N. Gkikas, *Automotive Ergonomics: Driver-Vehicle Interaction*, CRC Press, Boca Raton, Fla, USA, 2012.

[182] V. D. Bhise, *Ergonomics in the Automotive Design Process*, CRC Press, Boca Raton, Fla, USA, 2012.

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2013, Article ID 972570, 14 pages
http://dx.doi.org/10.1155/2013/972570

Research Article

The Front Seat Passenger: How to Transfer Qualitative Findings into Design

Sebastian Osswald,[1] **Petra Sundström,**[1,2] **and Manfred Tscheligi**[1]

[1] *CD Laboratory, ICT&S Center, University of Salzburg, 5020 Salzburg, Austria*
[2] *Mobile Life Centre, 164 29 Kista, Sweden*

Correspondence should be addressed to Sebastian Osswald; sebastian.osswald@gmail.com

Received 15 October 2012; Revised 4 February 2013; Accepted 26 May 2013

Academic Editor: Paul A. Green

Copyright © 2013 Sebastian Osswald et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

While the field of automotive HCI research abounds in driver-focused design and studies of driver distraction, relatively little attention has been directed to the front seat passengers. A challenge for designers and researchers is the ideation through exploration of passenger experiences and needs and the question of how to integrate them in design solutions. In this paper, we ground an ideation exercise based on results of a probing study conducted at two petrol stations. A probing package for front seat passengers was designed and distributed in cooperation with a petrol station company. Approximately 700 customers were approached randomly and 90 probe packages were handed out. 30 probing packages were returned and a segment clustering and a qualitative analysis were performed. The results gave insights into the activities and needs of front seat passengers and were then used in a design workshop together with a group of industrial designers. A set of ideas for novel interface solutions for front seat passengers were developed such as *the invisible engine* (look through the motor block with a display-based camera system). We further discuss the challenges and shortcomings of how to interpret and express ideas when transferring qualitative research findings into design.

1. Introduction

A great deal of time is spent driving—for business (to go to work and to appointments), for leisure (to visit family, friends, or a vacation destination), and for other purposes. The number of people within the vehicle changes, depending on the purpose for the trip. The driver may be alone or accompanied by passengers. For the front seat passenger, travel time can be a time to relax, spend family time together with those in the vehicle, or to complete tasks such as using Facebook, making phone calls, browsing the web, or watching a movie. However, the front seat passengers seem to be slightly forgotten in the field of automotive research. In most studies, the role of the front seat passenger is related to the task of driving the car, for example, assisting the driver with navigation (as in, e.g., [1]).

Every researcher and developer acknowledges that driving safely is crucial and designs and implements technology accordingly. Scientific research focuses on how to measure distraction in order to increase safety and on how to avoid situations that might affect safety (e.g., [2, 3]). The front seat passenger is most often regarded as additional support to the driver and the driving situation. However, drivers primarily travel in familiar areas (as reported in [4]) and drivers tend to use navigation systems when going to familiar destinations as reported by Lo and colleagues [5]. Thus, there are moments and situations when it is possible for the front seat passenger to pursue his/her own tasks. The questions here are as follows: What are the desires and needs of the passenger? How can we design for those desires and needs?

A more detailed analysis of how the social situation and the interaction in the car look like is desirable [6] in order to understand the relationship between driver and passengers and how the needs of front seat passengers differ from existing knowledge. Examining the car space itself should not mean to only focus on what concerns the driver; it should also involve everyone who is present in a car. In the car space, the driver is not an isolated individual, which is why the driver and the driving task are strongly affected by the presence of passengers and the social relationship between them. Further, safety concerns appear as passengers might distract the driver or the front seat passenger may interfere in the driving task. According to McEvoy and colleagues, the likelihood of a crash is almost 60% higher in the presence of two or more passengers in comparison to driving alone; distraction through talking to passengers was one of the most common distracting activities (29.8%) [7]. On the other hand, passengers could support the driving task with useful hints, for example, speed limits or through entering a destination in a navigation system. To make use of this potential, we are targeting the front seat passenger to explore and better understand the social space of the vehicle and derive inspiration for novel interface solution from passenger experiences.

In this paper, we present the procedure and results of a probing study that explores the desires and needs of front seat passengers. Cultural probing is a method used to inspire designers and engineers when designing interactive systems. The probing packages should support the participants to document and record specific feelings and experiences that can be used to stimulate imagination for novel interface solutions [8]. The probing packages were given to participants and included artefacts that serve as inspiration for the participants to gather data about their experiences and thoughts.

To draw inspiration from the results of the probing study, we presented the materials to a group of industrial designers. As stimulation and guidance to those interested in this design space, we provide descriptions of the most exciting and promising ideas we have collected. Within this paper, we discuss the issues and difficulties of presenting qualitative data to designers. There are two foci: (1) the step-by-step procedure of gaining inspiration for design ideas and (2) how to apply this procedure in a way that it tackles the challenges of transferring qualitative research findings into design.

We present the related work in the Section 2, followed by Section 3 which reports the topic of front seat passenger and application of cultural probing in the vehicle context. Within Section 3, probing materials, participants, procedure, and the results and data analysis are described. A detailed description of our design workshop is provided in Section 4, including the results according to a categorization provided by the participating researchers after the workshop. Section 5 contains the conclusion of transferring qualitative findings into design and the summary.

2. Related Work

Since Norman wrote his article saying how *"there's an automobile in HCI's future"* [9] in 2003, automotive research has vastly progressed in various areas. Norman feared *"that all the errors of the past, errors in nuclear-power control rooms, in process control rooms, in the control of ships, and commercial aviation will simply get repeated"* [9]. He was apprehensive that he or his colleagues would not be consulted when the engineering decisions in the field of automation are to be made. What we can see is that automotive research results find their way more and more into the automotive development process, for example, the rising interest from industry in conferences such as AutomotiveUI (Automotive User Interfaces and Interactive Vehicular Applications Conference (http://www.auto-ui.org/)). Nevertheless, we share the legitimate concern regarding fear that errors from the past will be repeated and argue for addressing the needs of every passenger in the design space of a vehicle to avoid distraction in the first place.

The importance of the car as a design space (e.g., [10–12]) has already been identified, but automotive HCI research has mainly been focusing on the driver, while often neglecting the other passengers in the car. A broader perspective of the social situation in the car was taken by Juhlin [13] in the field of automotive research. From an ethnographic approach on the social experiences of driving and travelling by car, they designed the Sound Pryer application [14] which is a "collaborative" MP3 player where the users (driver and/or passengers) can play their own music but also hear what other cars nearby are listening to. Brunnberg and Juhlin have also designed a backseat gaming platform [15] in which the world around the car is turned into a fantasyland filled with virtual creatures, treasures, and adventures. By pointing a gaming device towards objects they pass, the playing children can defend themselves against attacking creatures, pick up magic artefacts, or collaborate with players in traffic. Further challenges of ethnographical studies were discussed in the workshop "We are not there yet: enhancing the routine drives experience in the car context" [16] at the 2012 AutomotiveUI conference, focusing on statistical information about driving patterns and automotive technology trends "in order to make the boring and mundane attractive, entertaining, and engaging." Gridling and colleagues [17] furthermore conducted an ethnographic study in the car context itself, examining the collaboration between passengers and drivers. During various car trips, a researcher documented collaborative support, remarkable driving situations, and additional information that aimed to inform future in-car assistance. Other work conducted from a broader perspective on driving is the probing study by Wilfinger and colleagues [18], in which probing materials were given to children travelling in the backseat of cars. A result of this study was the RiddleRide game, in which questions about objects in the surroundings are posed at varying levels of difficulty to the people travelling together.

A challenge remains in that transfer of findings from ethnographic or qualitative research, such as a probing study, into design. The question is how to collect, formulate, and present the results of a qualitative study and transfer those results into a design process that inspires and guides design.

> *"In that process, the message (e.g., the qualitative user study results) is often skewed when passed on from one phase to another. Moreover, the way insights are communicated is often not inspiring enough. There is a lack of appropriate strategies and tools for how to communicate study results between different phases and between different stakeholders, so that informed design decisions can be made and data can be successfully used as*

> *inspiration for design tasks throughout the whole design process." [19]*

Therefore, this paper is not simply about a probing study we have performed, but also a solution of how we have presented the results of this study to a group of industrial designers to get inspiring new solution for interfaces for the front seat passenger. We decided to follow this procedure to find ways to transfer what we learn from users' practices into design, as this is something that is not easily done. In fact, this is one of the major "wicked" [20] problems of interaction design, the problem of transferring or translating qualitative user study findings into something that can be used in design. For supportive tools, we thus need richer descriptions of work that can act as case studies in this direction. The presented study and the transfer process document how self-reported findings can be used in design for future vehicle interfaces, the results that were obtained from the designing sessions, and the problems that were encountered in achieving these results.

3. Understanding Front Seat Passengers: The Probing Study

To open up the design space and to gather information on how front seat passengers would benefit from an IVIS, we handed out probing packages inspired by Gaver's cultural probes [8] to randomly selected front seat passengers approached at two locations in Salzburg, Austriain August 2011. The material was designed for application on longer car trips. The study duration was one month to ensure that the participants were together on at least one longer trip during this period. Longer car trips were well suited for probing materials that suggest, for example, drawing a novel interface, as a longer period of travel would be needed to complete the task. It was also possible to finish a longer task over the course of several short trips or to use a short trip to fulfill a quick task, such as noting impressions of a certain driving situation or describing specific experiences.

As opposed to diary studies, cultural probes provide a more intense way of gathering information and material about people and their thoughts and activities. Unlike traditional methods (e.g., usability testing, observation, and field studies), the technique separates the researcher from the process of gathering information and allows the participants to self-report. In user studies, the presence of the researchers often influences the behaviour and decision-making process, as the participants feel observed and remain in a socially controlled environment. Different results may occur if participants would be alone. Instructing the participants to report and reflect on situations that are particularly chosen and framed by a proper research agenda can help to decrease the negative influence emanating from the presence of researchers. This becomes especially important regarding the context in which a study is conducted. An automotive vehicle is a limited physical space for approximately five people. Placing a researcher in this context is to penetrate a private space, where most often only friends, colleagues, and relatives of a participant have access. It, therefore, seems valuable to give participants the opportunity to report by themselves about their experiences while driving, in order not to disturb the driver-passenger relationship and to gain further authentic results. Cultural probing as a qualitative method addresses these issue and provides a procedure for researchers to design materials that target at supporting the front seat passenger to report about certain driving situations.

The gathered results from cultural probes are particularly useful early in the design process to gain ideas and be inspired and for a further ideation phase. During the process, randomly selected participants are approached and invited, given a kit of materials, and briefed about the requirement to record or note specific events, feelings, experiences, or interactions over a set period. At the end of the in advance specified period, the materials are collected and analysed. In recent probing studies, the method of adding interviews in advance was introduced in order to supplement, validate, and explore the information gathered by the participants. The information gathered is then analysed and documented by the researchers.

3.1. Designing the Probing Materials. Originally, Gaver's probes [8] consisted of a diary, index cards, maps, a photo album/a media diary, and a disposable camera. Due to the safety-critical car context, we decided to alter the probes to be more suitable for front seat use. We reviewed the original probes materials according to their properties (what they are made of and how can they be used) and excluded the disposal camera, as we feared that the flash could distract or blind the driver. We tested a camera without flash in the car during different light conditions, but the results were not satisfying as the inside of the car was overall too dark. The geographic maps were replaced by a creative booklet, since the maps were used context dependently in the original probing package and not applicable to our study. Our booklet contained a set of context-related tasks and questions that aimed to make the front seat passengers reflect on their experiences and their role in the car. We wanted the passengers to stay within the context of a vehicle when working with the probes, in order to maintain the experience of being a front seat passenger inside a vehicle. The probes should remain with the car to enable the front seat passenger to document experiences every time directly in the vehicle context. Collaboration with other present passengers was encouraged, including the driver when possible. The various probing tasks were also possible to perform when the car had stopped moving.

Each of the cultural probe packages included a booklet, a 105 mm by 148 mm sheet of white paper, an envelope, paper-based 3D glasses, and 3–35 mm × 40 mm blank cards. In addition to the category-related sheets, the booklet contained a postcard and a treemap.

Using the provided tasks, we wanted the participants to think about their role as the front seat passenger, what they do while being the passenger, and how they behave while travelling by car. Passengers were encouraged to think about what is missing and what experiences they would like to have in the passenger seat. In total, ten inspirational situations were designed, divided into five categories: *good versus bad front seat passenger cockpit*, *information desire*, *information* *sharing*, *visions for a front seat passenger cockpit*, and *front seat passenger needs*.

For the category *good versus bad front seat passenger cockpit*, the participants were asked to judge the front seat passenger area and to describe the positive and negative aspects. The second inspirational idea was to write a short speech, accepting the Design Oscar for the best design of a front seat passenger device.

For the *information desire* category, we included paper-based 3D glasses with red/green glasses for the passengers, to put on and imagine the car being a 3D theater. We wanted to know what kind of augmented 3D events they could imagine in and outside the car. We then asked participants to write down what information they would like to have in a set of given situations (e.g., when it is raining or when backseat passengers are troublesome) in order to broaden our perspective on what information passengers may desire within the car.

Information sharing is an upcoming topic in the automotive domain [21] and was provocatively addressed by asking the passengers to write down travel related items they never would tell the driver on three little cards. A sealable envelope was added for the passengers to put their secrets inside. To encourage participants to think about a potential design, they were also asked to think about what applications (e.g., TV, games, or email) would be suitable for the driver, the passenger(s), and/or both.

Passengers were to be creative for the theme *visions for a front seat passenger cockpit*. We asked them to draw an IVIS and place it where they would like it to be within the car (e.g., in the car ceiling, on the door, or on the console). Passengers were also asked to use their imagination and draw futuristic technologies they would like to use in the front seat passenger area.

In the last theme, the *front seat passenger needs* were written on postcards so that the participants could imagine sending to a friend to complain about items they felt were missing regarding their regular activity as a front seat passenger. Table 1 gives an overview about all categories and materials used.

The strategy we took in designing our probes (see Figure 1) uses a provocative approach in which tactics of personal involvement are used to strip away participants' objections and open up for new perspectives [22].

3.2. *The Participants.*

As the probes method is used for a subjective view beyond what is graspable with methods like ethnography, we wanted to keep participants within the personal context of the car and chose not to bring them into the lab for a formal introduction seminar. We searched for an approach to recruit potential participants in context. Among several possibilities, we choose petrol stations as our recruiting area [23]. Petrol stations have a high frequency of visitors (possible participants) and passengers would have time, during the fuelling process, to speak with researchers. We were inspired by Kern and Schmidt's survey study that was conducted at a petrol station [24]. Due to the high frequency of potential participants arriving at the petrol station, we were not able to keep track of the exact number of passengers approached. Participants were selected randomly and approached regardless of gender or car type. To choose whom to approach of the many customers, quick decisions were made based on the license plate and whether a car was full with traveling accessories to distinguish between local customers and those simply passing through. Those with local license plates or from nearby cities were chosen to increase the chance of the participant's ability to return the probes to the petrol station. Participants were asked to return the probing materials to the petrol station where the materials were handed out. As Salzburg is close to the boarder of Germany, this included cities from Germany as there are a large number of people who drive to Austria to fuel their cars.

Figure 1: The probes: creative, descriptive materials.

Figure 2: A researcher distributing probing packages at a petrol station.

Having the participants return their probing package to the petrol station enabled them to independently return the materials and take their time to complete them and further does not require additional actions for shipping them. The collection of the materials was prearranged with the station staff. We expected a higher returning rate than in usual probing studies, as the packages could conveniently be returned upon the next visit to the petrol station. In retrieving and returning the packages at the petrol station, passengers were able to stay in vehicle context from the beginning to the end of the study. Figure 2 shows a researcher explaining the probing study to a petrol station customer.

3.3. *The Procedure.*

Drivers were often present when describing the goals and aim of the study to passengers. We used this opportunity to ask the driver to return the probes to the petrol station within a month's time, whereupon the driver would receive a voucher for a premium car wash as an incentive.

Table 1: Overview about the probing materials.

Category	Focus	Material
Introduction	A summary of the introduction to the study given by the researcher, providing information about the study procedure	2 pages in the booklet
Good versus bad front seat passenger cockpit	To write a 5-sentence assessment of the negative/positive aspects of the interior design of the front seat passenger area	4 pages in the booklet
Information desire	Documentation of informational needs that appear during a trip, determining the potential of augmented reality in the car	4 pages in the booklet, for example, categories, 3D glasses
Information sharing	Determining which information is useful for the front seat passenger and which information should not be shared with the driver, using cards to note down details concerning the driver	6 pages in the booklet, Tree diagram, paper, pencil, envelop, three cards
Visions for a front seat passenger cockpit	Sketch a computer especially for a front seat passenger	2 pages in the booklet, paper, pencil and stickers, sketch of a vehicle interior to complete
Front seat passenger needs	Documentation of activities during a car ride, preparation of a post card to write about a special car trip	4 pages in the booklet, post card in the booklet
Information	Informing the participant about the contact person responsible for the study, providing an address to return the package personally	2 pages in the booklet

(a)

(b)

Figure 3: (a) A participant conducting the initial task of sketching a future user interfcae for front seat passenger (b) and showing where in the car it should be placed.

Drivers were asked to do this as it was more likely that the driver would motivate the passenger to complete the probing tasks, than it would be for the passenger to motivate the driver to drive back to the petrol station. The incentive for the front seat passenger was thus indirect, as the completion of the task was to do a favour to the driver.

To introduce our study and the probes procedure and initiate the start of the study, we asked participants to draw a sketch for a future IVIS (see Figure 3(a)) while waiting at the petrol station. When finished, participants were asked to place the sketch in the best position within a car for their IVIS. The participants explained their sketch to the researcher, who additionally took a picture of the desired position of the sketched IVIS (see Figure 3(b)). Figure 4 depicts three images from the participants, one of which represents the rear seat and their children to express the need of being aware of their actions. The participant pointed out that it is annoying to turn around all time to observe the children and that a camera interface was desired. The second sketch expressed the need for an convenient integrated workplace in the front seat passenger area. The third sketch aimed at expressing the need for certain information to be combined into one location. Once the sketches were collected, participants received the probing booklet and probes and were told the study should be conducted while being a front seat passenger.

On average, it took about fifteen minutes for the explanation of the study and for participants to draw their IVIS sketch. Our general experience was that the approached drivers and passengers were mostly willing to answer questions during this time. Passengers who took part seemed to be curious and interested in the topic. Nevertheless, some people refused straight away, claiming to be in a hurry.

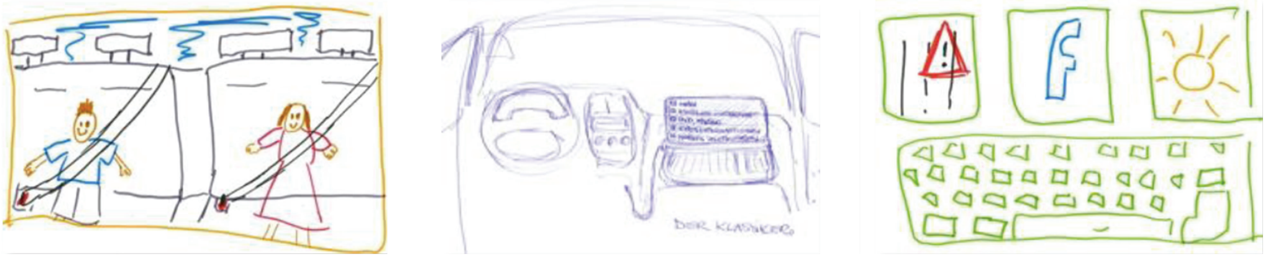


Figure 4: Three examples of sketches the passengers drew as the introductory task sketching for a future IVIS for the front seat passenger area.

3.4. The Results and Analysis. A total of 90 probing packages were distributed and 85 sketches were collected at the petrol station. (See Figure 4 for three examples of sketches.) Five sketches were missing as we allowed 10 participants who were in a hurry to complete the task later. Unfortunately, only five returned their IVIS sketches. After the preset period of one month, 30 probing packages were returned and collected from the petrol stations. All of these packages contained informational content and were usable for the qualitative analysis. As the probing was anonymous, no demographic data was collected during the study. We can only present the researchers subjective observation, where it can be said that the participants' age varied widely and that there was no tendency towards a single gender. It was noticed that two children participated and the rest of the participants appeared to be of legal age (older than 18 years).

Items that were left blank or contained notes about why a task was too difficult or considered not appropriate to complete were excluded. In total, 65 pages were left blank, whereby the drawing task from the category visions for a front seat passenger cockpit was left out most often. It is assumed that the initial drawing task at the petrol station affected the willingness of the participants to complete another drawing task. Looking through the rest of the materials, we received valuable feedback for every explored category. The category named *information desire*, in which the passengers were asked to use 3D glasses to imagine what additional things they could see in the car, received most responses. Every completed probing package contained this feedback. Ideas described included augmented overlays on buildings and landmarks presenting educational information about the surroundings, among others. Another wish was for more navigational information like a huge arrow that guides you through the traffic or a highlighting system for traffic signs, traffic lights, or safety-related events. The system would encircle, for example, the traffic signs with a glowing ring to make them appear more prominent. On the other hand, two participants mentioned that they would refuse any virtual overlay of information. The probes out of the category *visions for a front seat passenger cockpit*, in which the passengers had to sketch their future device, could also be described as successful. The system sketches showed the need for entertainment (TV/DVD/Movies), communication (social networking and video chat), or work support (email and organizer).

The 100 photos of the initial IVIS sketches placed in the vehicle were analysed in terms of placement. Figure 5 shows the distribution of the different positions mentioned. The most frequent placement was the central position of the glove compartment (35). The second most often position was on the visor (19). It is notable that the central right position in the windshield and the central right and lower right positions of the glove box were never chosen. Further, the driving mirror and the central console were not targeted, as they might be seen as areas intended solely for the driver.

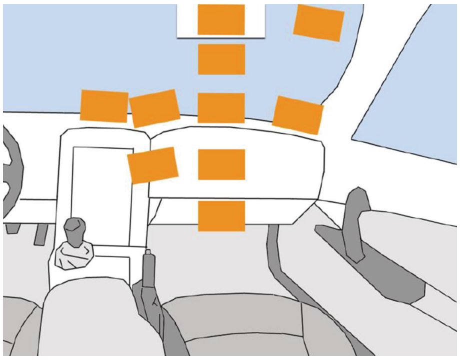


Figure 5: Overview about where participants placed the desgin sketches.

Most participants admittedly commented on how they were bad at drawing. However, we received 85 individual sketches. Regarding these sketches, we experienced that asking for certain features and for how these features shall be placed and related to each other was a viable method in the limited physical space of the passenger area in the cars. While we expected more general answers to the categories *front seat passenger needs*, *good versus bad front seat passenger cockpit*, and *information desire*, we were surprised to receive significantly more specific and personal answers such as *"I want information about if the next roadhouse provides wholefood"* and *"I need entertainment in the car and I need assistance to avoid the marital crisis every time I am driving with my husband. This is pure horror!"*

In order to better understand the strong and differentiated views of what is important in the vehicle, we counted and categorized all single answers and drawings. Even though we are aware of how this is far from the original idea of the probes [8], we noted how the same item, feature, or experience repeatedly appeared and became intrigued about how this will affect the design of a future interface. This allowed

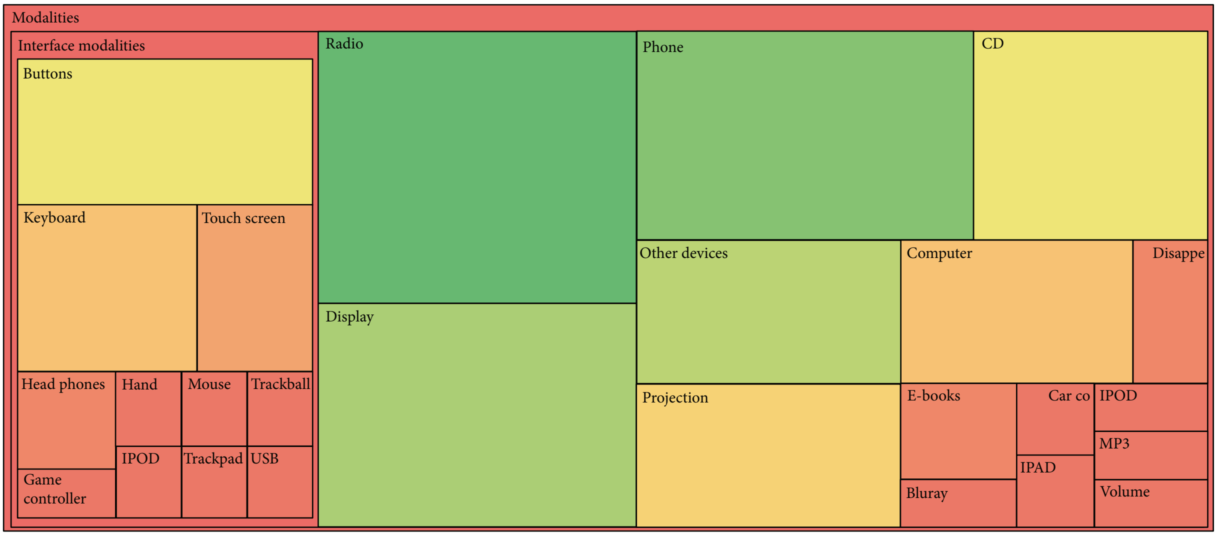


Figure 6: Segment clustering in NVIVO showing the distribution of the modality category; the size of each element correlates with how often an item was mentioned.

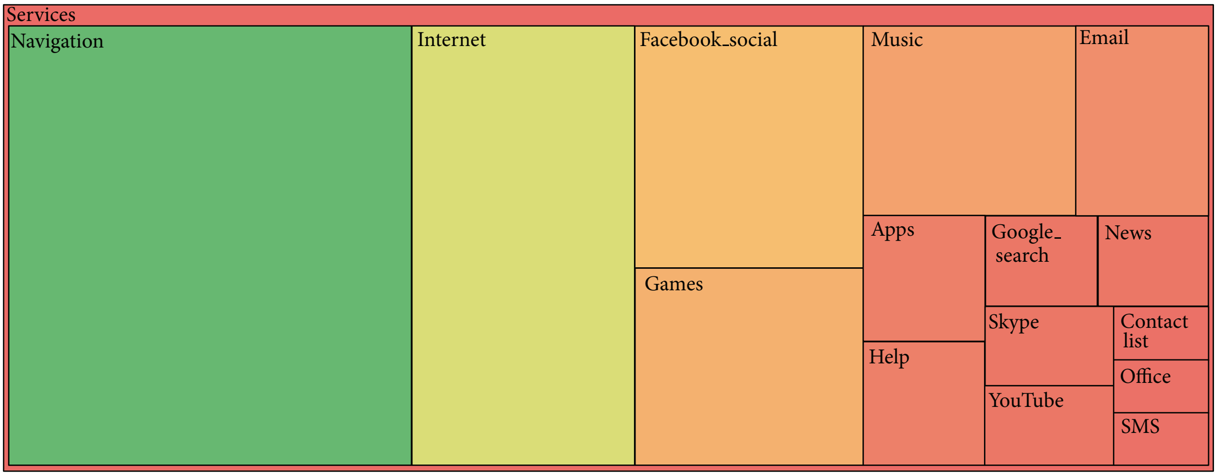


Figure 7: Segment clustering of the service category.

us to become familiar with the probes and to identify the relationships between the materials. What will be presented later in this paper is how the analysis of the returned probing materials in fact turned out to be an additional tool for the designers in a way we had not expected.

Summarizing the petrol station study, two researchers approached approximately 700 customers over one week (8 h a day) duration, distributed 90 packages, received 85 sketches, and took 100 pictures of the sketch placement. We received 30 packages with 235 filled out pages after a period of one month.

To ease the counting process and to build up more of an overall picture of the materials, we used the NVIVO software (http://www.qsrinternational.com/). This software allowed us to work with the unstructured materials and create cluster segments, referring to the occurrence and variety of the different requests that appeared in the materials. From the clustering, we determined four dominating categories of items requested: *modalities*, *services*, *context*, and *information*. The different sizes of the squares and the different colors in the treemap represent how often each item was mentioned in relation to the other requests.

Figure 6 shows the clustered treemap for the category modalities. According to the transcription of the probing materials, radio, display, and phone were mentioned most often in the modalities category, whereas we also see larger tiles for a projection interface and buttons. For the category services, Figure 7 shows that navigation was the most frequently mentioned functionality. The items internet, social services, and games followed closely. In Figure 8, the category context is presented. The most often referred contextual information was weather, followed by time and speed. Areas were also considered in contextual reference, which is why windshield and the backseat are presented in this overview. The category information is presented in Figure 9. The TV was most frequently mentioned, followed by DVD/movie and information about the surroundings of the vehicle. Items

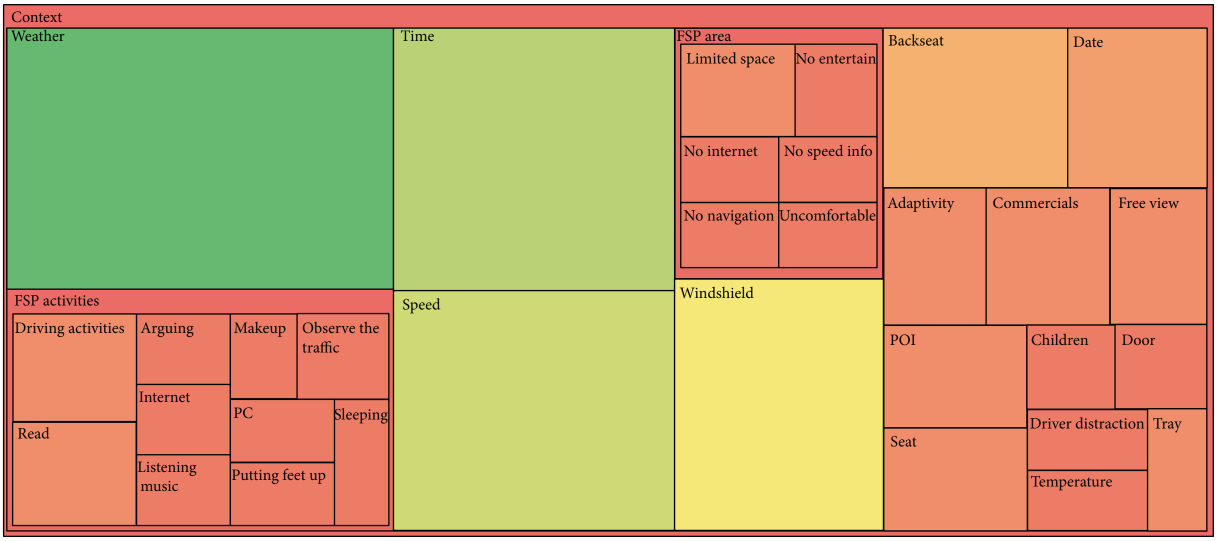


Figure 8: Segment clustering of the context category.

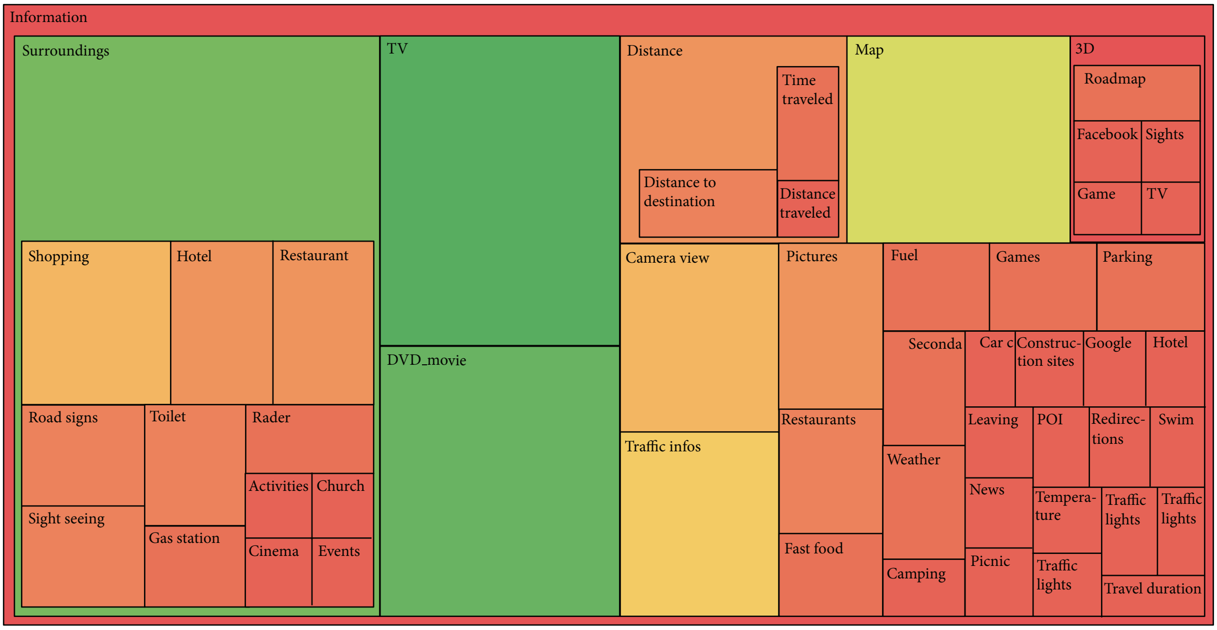


Figure 9: Segment clustering of the information category.

concerning a car trip like traffic, distance, or a map were also referred to.

This quantitative approach is uncommon when analysing cultural probing materials. The procedure of quantifying the data was similar to establishing an affinity diagram, a method commonly used within project management that allows large numbers of ideas stemming from brainstorming to be sorted in groups with focus on their relationships [25]. We used this approach to create an overview about what people think when they are engaged in discovering their personal space in the car. For us, it was useful to organize the content of the probing materials and to become immersed in the data. For the designers included later in the design process, the clusters highlight certain connection and idea categories. Summarizing the petrol station probing, the material revealed strong personal and emotional sequences that happen while driving (e.g., as the front seat passengers described their feelings about the driver in the envelop probe). During our analysis, the probes already inspired the participating researchers and discussions started about future automotive information systems.

4. Ideation: The Design Workshop

To investigate the value of the returned probes, we designed and organized a workshop where the whole material from the probing study was used as inspiration for a group of industrial designers in their sketching exercises for what could be a future IVIS for front seat passengers.

The five-hour workshop aimed to establish that generating valuable design ideas could be reached within an appropriate amount of time. To give the designers a chance of acquiring an understanding of the material in such little time, we used the segment clustering as a starting point to the material as we assumed that the clusters with its quantitative perspective would give the designers easier access to the materials.

4.1. The Participants. We invited five industrial designers from the University of Art and Design in Linz, Austria, to take part in this workshop. Two researchers (who were present in the probing study) participated in a workshop to introduce the topic and to support the designers. Industrial design, as a process, begins with the creation and development of a set of concepts and specifications that optimize the appearance, functions, and values for a system that aims to benefit the user and the manufacturer. Industrial designers are accustomed to developing concepts by considering the materiality of objects as well as the requirements of manufacturers and/or clients. This and the ability to place emphasis on aspects of a system that relates to characteristics of users' interests and needs grounded our decision to focus on industrial design before continuing into the processes of interaction design.

Each of the designers received an incentive of €50 for the time they spent in the workshop and for the work they completed during the session.

4.2. The Procedure. To build on the designers' skills to create ideas out of inspirational sources (in this case the probes in their original form in combination with the clustering), we decided to apply the nominal group technique (NGT) and combine it with a speeded sketching approach [26], as this method already showed its potential for designers to gain ideas. This procedure was chosen to allow for a session where the designers would get inspired and sketch a set of ideas in a brainstorming-like fashion. The ideation was thereby limited in time and to a set of certain tasks (four in total, referring to the clusters described in the results and analysis section of the probes information, modalities, services, and contextual information). The idea was that the restrictions would increase the attending designers' creativity [26].

The NGT was originally developed by Delbecq and Van de Ven [27]. This technique is a structured variation of a small-group discussion to reach consensus and encourage contribution from everyone involved. The process encourages both, individual thinking and group discussion, as it prevents the discussion or the work to be completed predominantly by a single person alone. In the original NGT process, a small group of five to six members (experts) are seated around a table and introduced to a topic. They are allowed to browse through supplied material to gather a holistic perspective. Thereafter, the researcher initiates an initial brainstorming session with an open-ended question in the area of interest. Each expert spends several minutes in silence individually brainstorming, noting as many ideas as possible. Each idea is then introduced to the group, where no criticism is allowed, and only questions for clarification could be posed. Next, the ideas are discussed to clarify meanings, explain intended logic, and state agreements or disagreements. In the end, the ideas are prioritized using a multivoting or list reduction process. The overview about this process is documented in Table 2.

Table 2: Overview about the NGT inspired workshop procedure.

Step	Task
1	A small group of 5 or 6 members (experts) are seated around a table.
2	The collected probing materials are presented to the group, it is allowed to browse through the whole material.
3	An open-ended question is stated: In which way could we redesign the car cockpits that it fit to the imaginations of potential drivers?
4	Each expert spends several minutes in silence individually brainstorming and notes down one idea.
5	Each expert sketches the noted idea and adds two additional explaining sentences.
6	The ideas and sketches are presented to the group, no criticism is allowed, but clarification in response to questions is allowed.
7	All ideas are discussed after presentation to clarify meanings, explain their logic and to state agreements or disagreements.
8	The best ideas are selected and the votes are shared within the group.
9	A final discussion about the best idea is encouraged.

For our purposes, we separated the workshop into five main stages. First, a brief warm-up and introduction of the probes and the study is presented. Then, designers could familiarize themselves with the supplied materials (the booklets, the in-situ sketches, and the segment clustering). Similar to the NGT process, we then allowed for an individual brainstorming session, reflecting on the task and sketching ways to approach it. We followed this by having a presentation session to discuss and evaluate these ideas as a group. The last stage included a discussion on the procedure and materials of the task.

4.3. The Results: Ideas That Were Generated. In total, the designers generated 45 sketches. Of these sketches, we chose idea categories to summarize and classify the level of innovation of the ideas. The classification was done in an expert evaluation session with the goal to estimate the level of innovation for the ideas. Three HCI experts, with backgrounds in automotive user interface research, participated and rated the ideas according to the following three categories. First, *state-of-the-art-level* ideas represent approaches that could be included in available vehicles on the market. The second category, *ideas with potential*, summarizes ideas that were connected to available technology that may be realized in

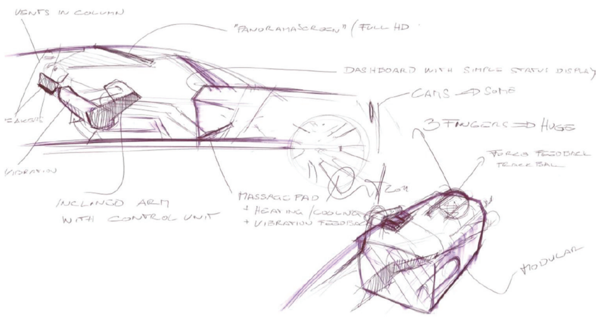


Figure 10: Armrest interface: thumb-controlled device and roof top interface.

the near future. *A future outlook* focused on ideas that explore novel approaches of experiencing the vehicle and stand out in terms of unique attributes. Each of these categories contains the original ideas designed in the sketching session in more detail, listed in ascending order from state-of-art level to more innovative solutions. All of the ideas were derived from the probing materials presented. We asked the designers to determine a certain material for each sketch, but we experienced that the majority of ideas were generated based on multiple impressions, making it difficult to assign a specific material to an idea. Nevertheless, some materials were described as more inspiring, which will be discussed in the following section. Safety aspects were also widely discussed, as the ideas violate laws, standards, and design guidelines, that need to be considered when developing information systems for the automotive market. We decided to neglect safety issues in this stage of ideation, as we felt this might reduce the attractiveness of the ideas and reduce the creativity of the designers. However, we do acknowledge that safety aspects need to be addressed if any of these ideas are further developed.

4.3.1. State-of-the-Art Level. Most of the sketches present the various *modalities* that were mentioned in the probing, such as touch interaction, buttons, and displays in an integrated form. We tried to divide the materials between the designers but almost all of them ended up designing for a modality centric perspective. A frequent suggestion was a basic control unit in the car, where the passenger could dock his/her mobile phone or smart phone and gain access to larger screens and more functionality in the car area. This would benefit from Internet access provided by the vehicle's communication platform and further allow storing vehicle-related settings on the personal phone device. One specific idea in this direction was a docking station in the door area that would offer directional touch interaction to enable the ability to easily browse through listed content coming from both the phone and from services stored in the car. This would be possible through an ergonomically designed thumb device positioned where the traditional armrest now is (see Figure 10).

Another idea was a concept of overlaying information in an IVIS, positioned in the central console, where it would display different content that would only be viewable from certain angles, allowing the passenger a private interaction space without disturbing the driver. The idea was to create a *capsule* within which the passenger gets the possibility to forget about the surroundings and relax. It was discussed how a separation might affect the social relation between a driver and a front seat passenger. The idea was, however, based on an even more social action described in our probes. Some front seat passengers described that they often try to sleep while they are on the front seat in situations when they want to recover from a previous driving period on a long distance drive. The *capsule* addresses this issue and tries to establish an environment for the front seat passenger to relax. In this direction, the designers also suggested various alternatives for private speakers and displays, for example, in the head rest or rooftop. This would allow the passenger and the driver to hear/see their own music/media in the car without disturbing each other.

Another designer had a well-being perspective and pointed out that the essential feature of the car is the "*aesthetics of the surroundings*." He sketched the front seat passenger area in a way in which the elements were to be nonintrusive through reductions of the elements' visibility. He reported that he deliberately included any feature he could think of (like a head-up display, a keyboard, and a touch screen), to envision a multipurpose cockpit for the front seat passenger (see Figure 11).

The window on the right hand side of the passenger was mentioned in connection with an interactive screen that would reserve a certain area for the driver to see the outside mirror on that side. This screen would enable the passenger to look out and see various augmented content in the passing environment.

Several of the designers' ideas addressed the question of how the entertainment desires of the passengers should be fulfilled. Various alternatives for using the windshield, for both 2D and 3D movies, were suggested. Issues related to driver distraction were heavily discussed and sketched in various ways (e.g., avoiding display reflections through altering the position or applying visual filters in head mounted devices to handle driver distraction caused by motion of projected objects in the windshield).

Having in mind that the designers started from scratch, these ideas provide an initial impression of the potential that this area has. The ideas in this section are very close to many state-of-the-art solutions, such as the overlaying information concept, which is very similar to the SPLITVIEW system from Daimler (http://www.daimler.com/). There are similar concepts and systems available on the market and in the literature that can be closely matched to ideas presented in this section.

4.3.2. Ideas with Potential. A rather simple but effective idea that was suggested in several ways was to remove the glove compartment and replace it with either a screen that could slide out of the console in front of the passenger or a table-like area for a keyboard and other modalities to navigate a docked phone device with. Similar to this slide out/in solution,

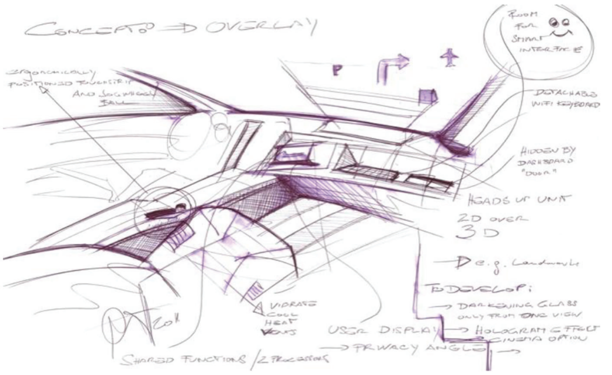


FIGURE 11: Aesthetic overlay concept with reduced element visibility.

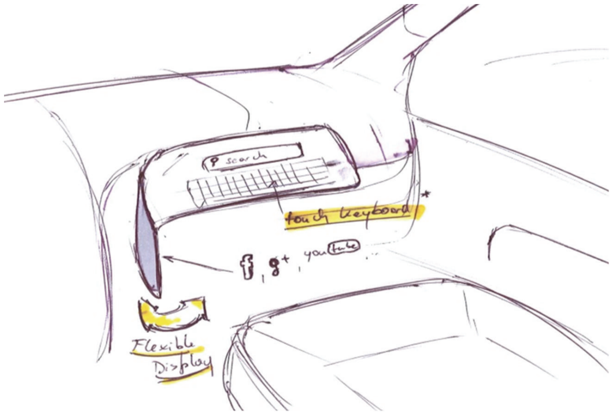


FIGURE 12: A flexible display, bendable for more privacy and without disturbing the driver.

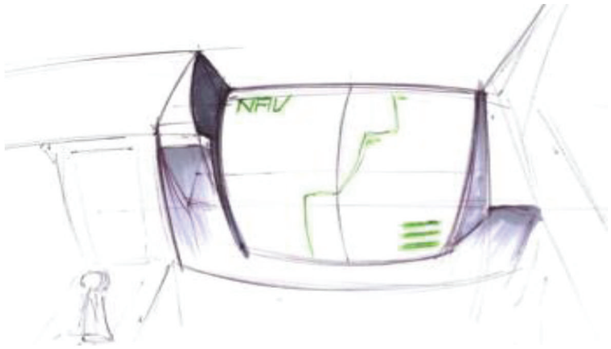


FIGURE 13: An integrated block interface physically separated from the console.

a rail-based system was proposed that would use the most often idle space in the ceiling of cars. This solution would also serve for additional modalities and a screen that could be pulled down in front of the passenger. Figure 12 shows a moving display that was derived from an idea, inspired by the usage of the envelope probes themselves. This display can be bent towards the passenger, leading to a more private space that is not visible for the driver, like the cards in the envelopes.

The *roof top interface* idea (as sketched in Figure 10) appeared in other ideas also. The most frequently suggested idea for usage of this space was to install a ceiling screen that the passenger could see by tilting the seat back and making use of free space in the rear seat area if no passenger is sitting there. The concept was to pull the passenger out of the context of the car and provide a private space for watching films or for reading.

One of the more advanced ideas on the concept of a docking station for a personal phone device was the idea of letting the passenger send messages to be displayed on the outside of the car, like status updates in Facebook. One suggestion for this was to place an AMOLED (http://www.techopedia.com/definition/24217/active-matrix-oled) display (a display technology commonly used in mobile devices) in the area between the wing mirrors and the tail light following the upper edge of the doors to let a message run from the front to the back. We do not see these docking ideas as remarkable, due to existing connections between the car and a personal phone device. These ideas consider basic services like handling phone calls and so on; however, more advanced services in this direction are not that far off.

4.3.3. A Future Outlook. A few of the sketches were more futuristic and innovative. One such idea contained the *gaming by foot* idea. This gaming device can be slid into the passenger's footwell to allow for gaming through foot interaction while riding in the car. We especially like this idea, as it uses an available area in the car and also holds great potential for an in-car game (e.g., imitating the driving style of the driver or silently interacting with the passengers in the backseat for a group-based game). Further, we like this idea because it is an unusual interaction modality that would not disturb the driver if designed correctly. The gaming activities could take place in an augmented form for the inside and the outside of the car.

Another innovative idea is *the invisible engine*. This idea uses a block interface, as presented in Figure 13. This interface is situated in the area of the glove compartment, but it is innovative in the design of the front seat area. In this idea, it was suggested that the information would *glow* out of the dashboard. The surface of the interface should be similar to the dashboard and only the presented information should be visible in fine green, pulsating lines. The area could also be used as a window connected to a camera in the front of the car, allowing for the experience of seeing through the engine, looking out of a window right above the road.

To conclude this section, we want to address the fact that, to some extent, the ideas presented here are affected by how we wanted to find possible passengers' needs when not having to assist the driver. This focus directed us slightly away from other kinds of social interactions the passenger could have had with the driver that would not make him or her the codriver. Furthermore, our continuous research aimed to include interaction that could take place between the front seat passenger and the potential passengers in the back. However, foci have been directed almost exclusively towards the driver and the needs of the driver and potentially towards the issues of entertaining some children in the back [18]. We believe a more directed focus on the front seat passengers alone is needed as an initial step towards a safe but more holistic and engaging driving experience for everyone in the car. After all, not everyone is driving as a family of four and not every driver needs help or wants to continually interact while driving.

4.4. A Reflective Discussion on the Provided Materials. After the design exercise, we had a discussion with the designers regarding the task itself, the provided materials, and the topic of design, in general, and industrial design, specifically. This discussion was recorded by one of the present researchers to allow for documenting the arguments presented below.

When asked for the most inspiring materials, the designers liked the sealed envelopes with the secret items the passengers did not want to tell the driver. They stated their belief that the passengers knew these envelopes were to be opened and read. However, designers thought these envelopes allowed the passengers to more freely express their feelings. One designer enjoyed the impression that these notes revealed the true emotions of the passengers, such as *"I do not like the smell of the driver"* or *"I'm really scared every time he is passing another car."* He found these notes entertaining and used them to gain new inspiration when he was needing it.

> *"I really liked the envelopes, because to me they represent emotional verbal statements, and they tell me more than the other stuff as they express emotions." (Designer 1)*

The designers expressed several times how they also liked the quantitative data. One of the designers explained how the quantitative data served as a validate tool to make sure that what he designed others would like too. Another designer said it was a way for him to see what the majority wanted. And when then asked if the unique and, perhaps, the more innovative requests excited him more, he stated how the car is a significantly more expensive device where individuality cannot be designed for in the same way as for less expensive items, for example, mobile phones. When further asked if these passengers requests could not be seen as additional equipment when purchasing a car, he answered:

> *"Yeah, but the additional stuff must not be added, like you glue it on or something, it needs to be integrated into the whole system." (Designer 4)*

Another designer phrased it as follows:

> *"You know, you do not want to have the same ride as everybody else, but the advantage of having the same ride as everyone else is that it's cheaper, and if you want an individual ride it costs more." (Designer 3)*

Further, the designers stated how they saw it as their task to take the requests of the users and break them down and combine the requests more into one integrated device. Referring to the various items and features that were asked for, one of the designers expressed it as follows:

> *"It is like this Swiss army knife coming out of the dashboard, but this represents how people think . . . from those parts I can take most of my conclusions and put them together with the drawings and the emotions." (Designer 1)*

Furthermore, the designers felt they had too little time for this task. However, they explained that in contrast to how they normally would need to spend a month's time to research on a topic, the provided materials were helpful in the same way in a much shorter time. They also stated that the topic itself also helped them since everyone is familiar with cars and they could empathize with being the front seat passenger, as well as the driver. They told us about a project where they were to design new ovens. This task had been harder for them as they were less familiar with ovens.

5. Conclusions

In this paper, we presented the results of a probing study conducted at two petrol stations. We aimed to explicitly identify the steps taken towards transferring the results of this study into the field of design. We focused on the front seat passenger instead of the driver, as we see potential in addressing the passengers' needs and experiences to improve travel experiences. Boehner and colleagues have already concluded that it appears in the literature, as if probe responses often lead directly to final system designs [19]. In the original probes work [8], there is an important intermediary step in interpretation called exploration and sketching. This is where the returned probes are interpreted and expressed through a series of *ideas* for different designs. As researchers and HCI experts without a deep background in conceptual design, we wanted to turn the results of our probing study over to a group of external designers. Transferring the quantitative data into design adds yet another important interpretational step to the multilayered process of expression and interpretation suggested by Gaver and colleagues [22]. Figure 14 shows a modified version of Gaver and colleagues' original procedure sketch, to which we have added this intermediary step of collecting and expressing qualitative user study results in a way that inspires design and design thinking.

In terms of the study procedure, the recruiting process in an area where the front seat passenger target group can be approached in-situ turned out to be successful compared with a previous conducted probing study [18] and further beyond the high number of creative materials that could be collected. Moreover, it proved to be successful to use a combined qualitative and quantitative approach in presenting the results to the group of experts—the industrial designers. Regarding the outcome of the workshop and the described ideas, it can be stated that this approach would be beneficial for anyone approaching a topic where the items worked with (i.e., cars) are slightly more expensive and when the requests of a majority have to be considered.

Some of the probing materials did not work as well as others (e.g., the sketches were returned in high numbers, but most of the passengers did not fill in the related pages of their own needs within the booklet). The workshop with

Figure 14: Our modified version of Gaver and colleagues' multilayered process of expressing and interpreting the probes [22].

the industrial designers could benefit from a longer timeframe (up to one month, depending on the targeted state of the ideas), to give more room for the designers' regular iterative procedures and perhaps to create more sophisticated ideas. From the perspective of the ideas themselves, they are to be considered as an initial step, although a valuable one, in the ideation phase and could serve as a starting point for anyone interested in the topic of designing for the needs and interests of front seat passengers. Our future work will address the challenge to iterate the probing method to be more suitable for the car context. We experienced drawbacks in some steps we took but managed to create valuable sources of inspirational material with our probing packages. A challenge in the future will be to develop probing packages that are standardized, to a certain degree, and adapted to the addressed target group, such as front seat passengers, drivers, or rear seat passengers. As described in the development process of our probes, the influence on safety due to the driving situation is a critical factor for probing studies that need to be considered regardless of the target group within the car.

Acknowledgments

The financial support by the Federal Ministry of Economy, Family and Youth and the National Foundation for Research, Technology and Development is gratefully acknowledged (Christian Doppler Laboratory Contextual Interfaces). The authors would like also to thank the Industrial Design Department of the University of Art and Design Linz.

References

[1] B. Brown and E. Laurier, "The normal natural troubles of driving with GPS," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, pp. 1621–1630, ACM, May 2012.

[2] S. Osswald, A. Meschtscherjakov, D. Wilfinger, and M. Tscheligi, "Interacting with the steering wheel: potential reductions in driver distraction," in *Proceedings of the International Joint Conference on Ambient Intelligence (AmI '11)*, pp. 11–20, Springer, November 2011.

[3] I.-M. Jonsson, C. Nass, J. Endo et al., "Don't blame me I am only the driver: impact of blame attribution on attitudes and attention to driving task," in *Proceedings of the Extended Abstracts on Human Factors in Computing Systems (CHI EA'04)*, pp. 1219–1222, ACM, 2004.

[4] G. Leshed, T. Velden, O. Rieger, B. Kot, and P. Sengers, "In-car GPS navigation: engagement with and disengagement from the environment," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*, pp. 1675–1684, Florence, Italy, April 2008.

[5] V. E. Lo, P. A. Green, and A. Franzblau, "Where do people drive? Navigation system use by typical drivers and auto experts," *Journal of Navigation*, vol. 64, no. 2, pp. 357–373, 2011.

[6] A. Meschtscherjakov, D. Wilfinger, N. Gridling, K. Neureiter, and M. Tscheligi, "Capture the car! qualitative in-situ methods to grasp the automotive context," in *Proceedings of the Automotive User Interfaces and Interactive Vehicular Applications (UI '11)*, pp. 105–112, ACM, 2011.

[7] S. P. McEvoy, M. R. Stevenson, and M. Woodward, "The prevalence of, and factors associated with, serious crashes involving a distracting activity," *Accident Analysis & Prevention*, vol. 39, no. 3, pp. 475–482, 2007.

[8] B. Gaver, T. Dunne, and E. Pacenti, "Design: cultural probes," *Interactions*, vol. 6, no. 1, pp. 21–29, 1999.

[9] D. A. Norman, "There's an automobile in hci's future," *Interactions*, vol. 12, pp. 45–54, 2005.

[10] D. J. Wheatley, "Beyond the desktop: and into your vehicle," in *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI EA '00)*, pp. 43–44, ACM, Hague, The Netherlands, 2000.

[11] A. Schmidt, A. K. Dey, A. L. Kun, and W. Spiessl, "Automotive user interfaces: human computer interaction in the car," in *Proceedings of the 28th Annual CHI Conference on Human Factors in Computing Systems (CHI EA '10)*, pp. 3177–3180, ACM, April 2010.

[12] D. Kern and A. Schmidt, "Design space for driver-based automotive user interfaces," in *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '09)*, pp. 3–10, ACM, 2009.

[13] O. Juhlin, "Social media on the road: mobile technologies and future traffic research," *IEEE Multimedia*, vol. 18, no. 1, pp. 8–10, 2011.

[14] M. Östergren and O. Juhlin, "Car drivers using sound pryer—field trials on shared music listening in traffic encounter," in *Consuming Music Together: Social and Collaborative Aspect of Music Consumption Technologies*, K. O'Hara and B. Brown, Eds., pp. 173–190, Springer, 2006.

[15] L. Brunnberg and O. Juhlin, "Movement and spatiality in a gaming Situation—boosting mobile computer games with the highway experience," in *Proceedings of the IFIP Conference on Human-Computer Interaction (INTERACT '03)*, pp. 407–414, IOS Press, 2003.

[16] C. Montesinos, D. Szostak, and A. Zafiroglu, "We are not there yet: enhancing the routine drives experience," in *Adjunct Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (UI '12)*, pp. 249–256, 2012.

[17] N. Gridling, A. Meschtscherjakov, and M. Tscheligi, ""I need help!" exploring collaboration in the car," in *Proceedings of the ACM Conference on Computer-Supported Cooperative Work (CSCW '12)*, pp. 87–90, Seattle, Wash, USA, February 2012.

[18] D. Wilfinger, A. Meschtscherjakov, M. Murer, S. Osswald, and M. Tscheligi, "Are we there yet? A probing study to inform design for the rear seat of family cars," in *Proceedings of the 13th IFIP TC13 Conference on Human-Computer Interaction (INTERACT '11)*, pp. 657–674, ACM, 2011.

[19] K. Boehner, J. Vertesi, P. Sengers, and P. Dourish, "How HCI interprets the probes," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '07)*, pp. 1077–1086, ACM, San Jose, Calif, USA, May 2007.

[20] R. Buchanan, "Wicked problems in design thinking," *Design Issues*, vol. 8, no. 2, pp. 5–21, 1992.

[21] O. Inbar and N. Tractinsky, “Make a trip an experience: sharing in-car information with passengers,” in *Proceedings of the 29th Annual CHI Conference on Human Factors in Computing Systems (CHI ’11)*, pp. 1243–1248, ACM, May 2011.

[22] W. W. Gaver, A. Boucher, S. Pennington, and B. Walker, “Cultural probes and the value of uncertainty,” *Interactions*, vol. 11, no. 5, pp. 53–56, 2004.

[23] S. Osswald and M. Tscheligi, “Gas station probing: the front seat passenger,” in *Proceedings of the Adjunct 3rd International Conference on Automotive User Interfaces and Vehicular Applications (UI ’11)*, 2011.

[24] D. Kern and A. Schmidt, “Gas station flash survey—a method for interviewing drivers,” in *Proceedings of the Workshop on Automotive User Interfaces. Mensch und Computer*, pp. 41–45, Bauhaus-Universität Weimar, 2007.

[25] H. Beyer and K. Holtzblatt, “Contextual design,” *Interactions*, vol. 6, no. 1, pp. 32–42, 1999.

[26] J. Wentzel and S. Holmlid, “Speed sketching with designers: user inspired brainstorming,” in *Proceedings of the Designing Pleasurable Products and Interfaces*, ACM, 2009.

[27] A. Delbecq and A. Van de Ven, “A group process model for problem identification and program planning,” *Journal of Applied Behavioral Science*, vol. 7, no. 4, pp. 466–492, 1971.

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2013, Article ID 951762, 8 pages
http://dx.doi.org/10.1155/2013/951762

Research Article

Human's Overtrust in and Overreliance on Advanced Driver Assistance Systems: A Theoretical Framework

Toshiyuki Inagaki and Makoto Itoh

Faculty of Engineering, Information and Systems, University of Tsukuba, 1-1-1 Tennodai, Tsukuba 305-8573, Japan

Correspondence should be addressed to Makoto Itoh; itoh.makoto.ge@u.tsukuba.ac.jp

Received 26 December 2012; Accepted 24 March 2013

Academic Editor: Klaus Bengler

Copyright © 2013 T. Inagaki and M. Itoh. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper gives a theoretical framework to describe, analyze, and evaluate the driver's overtrust in and overreliance on ADAS. Although "overtrust" and "overreliance" are often used as if they are synonyms, this paper differentiates the two notions rigorously. To this end, two aspects, (1) situation diagnostic aspect and (2) action selection aspect, are introduced. The first aspect is to describe overtrust, and it has three axes: (1-1) dimension of trust, (1-2) target object, and (1-3) chances of observation. The second aspect, (2), is to describe overreliance on the ADAS, and it has other three axes: (2-1) type of action selected, (2-2) benefits expected, and (2-3) time allowance for human intervention.

1. Introduction

Driving a car requires a continuous process of perception, cognition, action selection, and action implementation. Various functions are implemented in an advanced driver assistance system (ADAS) to assist a human to drive a car in a dynamic environment. Such functions, sometimes arranged in a multilayered manner, include (a) perception enhancement that helps the driver to perceive the traffic environment around his/her vehicle, (b) arousing attention of the driver to encourage paying attention to potential risks around his/her vehicle, (c) setting off a warning to encourage the driver to take a specific action to avoid an incident or accident, and (d) automatic safety control that is activated when the driver takes no action even after being warned or when the driver's control action seems to be insufficient [1]. The first two functions, (a) and (b), are to help the driver to understand the situation. Understanding of the current situation determines what action needs to be done [2]. Once situation diagnostic decision is made, action selection decision is usually straightforward, as has been suggested by recognition-primed decision making research [3]. However, the driver may sometimes feel difficulty in action selection decision. Function (c) is to help the driver in such a circumstance. Note that any ADAS that uses only the three functions, (a)–(c), is completely compatible with the *human-centered automation* principle [4] in which the human is assumed to have the final authority over the automation.

Suppose an ADAS contains the fourth function, (d). Then the ADAS may not always be fully compatible with the human-centered automation principle, because the system can implement an action that is not ordered by the driver explicitly. Some automatic safety control functions have been already implemented in the real world. Typical examples are seen in an advanced emergency braking system (AEBS) and a lane departure prevention system (LDP). When a vehicle is approaching to the forward vehicle, AEBS tightens the seat belt and adds a warning to urge the driver to put on the brake. When the system determines that the driver is late in braking, it applies the brake automatically based on its decision. LDP is an automatic system that applies the brakes to individual wheels, without any intervention of the driver, to prevent the vehicle from departing the lane. The fact that the driver may not always be kept as the final authority over the automation in such ADAS does not necessarily mean that those designs should be prohibited. On the contrary, the automatic safety control functions are effective and indispensable to attain driver safety, which suggests the domain dependence of human-centered automation [5]. It is true, however, that careful investigations are needed regarding to what extent the system may be given authority for deciding and acting autonomously without asking the human driver's approval or consent, because autonomy of smart machines sometimes brings negative effects, such as the out-of-the-loop performance problem, loss of situational awareness, complacency or overtrust, and automation surprises; see, for example, [6–10].

Moreover, as for the fourth function, (d), the following question is frequently asked: "when the ADAS is capable of coping with the situation automatically without any intervention of a driver, is not it possible for the driver to be overly reliant on the system and give up active involvement in driving?" For instance, the Ministry of Land, Infrastructure, and Transport as well as the National Police Agency of the Government of Japan has been somewhat discreet in introducing highly automatic safety control functions into ADAS on concern that the drivers may place "*overtrust*" in or "*overreliance*" on automation. However, discussions regarding overtrust and overreliance have not been rigorous enough yet until this point. As ADAS becomes smarter and more autonomous, these issues attract more serious concerns worldwide, for example, ASV project in Japan, HAVEit, and ISi-PADAS projects in EU.

Aviation domain has various studies regarding overreliance on automation; see, for example, [11–14]. Suppose that the automation is very rare to miss detections (i.e., it almost always alerts the human when an anomaly or an undesirable event occurs). Although a given alert is likely to be false, the human can be confident that there is no undesirable event as long as no alert is given. The human accordingly does not take precautions while the automation gives no alert. Meyer [13] has used the term *reliance* to express such a response of the human. If the human assumed that the automation will always give alerts when an undesired event occurs, that may be overtrust in the automation's capabilities, and the resulting reliance on the automation can be overreliance.

The relevant term, *complacency*, is usually defined as "self-satisfaction especially when accompanied by unawareness of actual dangers or deficiencies" [15]. However, the term is often used in human factors area to express a phenomenon that the human does not monitor the automation. Moray and Inagaki [16] have pointed out that the usage is misleading, because "not monitoring the automation" does not necessarily mean that the human is complacent. An obvious counterexample is that the human is busily occupied with extremely urgent tasks. Therefore, this paper tries to avoid using the term complacency.

This paper proposes a theoretical framework to describe, analyze, and evaluate the driver's overtrust in and overreliance on ADAS. Although the two notions, overtrust and overreliance, are often used as if they are synonyms, this paper differentiates the notions rigorously. To this end, two aspects, (1) situation diagnostic aspect and (2) action selection aspect, are introduced. The first aspect is to describe overtrust, and it has three axes: (1-1) dimension of trust, (1-2) target object, and (1-3) chances of observation. The second aspect, (2), is to describe overreliance on the ADAS, and it distinguishes other three axes: (2-1) type of action selected, (2-2) benefits expected, and (2-3) time allowance for human intervention.

2. Overtrust

Overtrust can be defined as a psychological state in which the human trust is inappropriately high. Overtrust is an incorrect situation diagnostic decision claiming that the object is trustworthy when it actually is not. This paper introduces three axes for describing the types of overtrust in a precise manner.

2.1. Dimension of Trust. The first axis (1-1) gives the dimension of trust. Lee and Moray [17] have distinguished four dimensions for trust: (a) *foundation*, representing the fundamental assumption of natural and social order, (b) *performance*, resting on the expectation of consistent, stable, and desirable performance or behavior, (c) *process*, depending on an understanding of the underlying qualities or characteristics that govern behavior, and (d) *purpose*, resting on the underlying motives or intents. Trust in an object is appropriate when all the dimensions are evaluated correctly. When there is a dimension that is evaluated inappropriately high, perceived trust is seen as overtrust. Therefore, some types can be distinguished for overtrust depending on which dimension of trust is violated.

Example 1. Suppose the driver thought that "the ADAS has been successful in coping with the situations so far. I am sure that the system will continue to be successful hereafter, too." This is a type of overtrust, violating the second dimension of trust.

Example 2. Imagine a case in which the driver thought that "I do not know how the function is implemented in the ADAS. I am not informed how the task is carried out, either. However, it would be quite alright even if I do not know the details." This is a type of overtrust, violating the third dimension of trust.

Example 3. Assume that the driver said that "I do not understand why the system is doing such a thing. However, the system should be doing what it thinks it necessary and appropriate. The system will not harm us." This type of overtrust does not satisfy the fourth dimension of trust.

Itoh [18] developed a model of human trust in automation and discussed the relationship between the three dimensions of trust, that is, purpose, process, and performance. The model takes into account the function of an automated system, limitation of the working conditions for the function, and the reliability of the automation function within the limitation. User's misunderstanding of the function is related to overtrust in terms of the purpose dimension. Expecting successful work of the automation beyond the limitation is a type of overtrust, violating the dimension of process. On the other hand, human's complete trust in automation within the limitation of the prescribed working conditions may not be overtrust if the reliability of the automation is perfect within the conditions. Itoh [18] also suggested that increase of trust in terms of performance may result in the overtrust in terms of process, and finally in the overtrust in terms of purpose. Such expansion of overtrust was called "ripple effect."

2.2. Target Object to Which Overtrust Is Addressed. The second axis (1-2) describes a target object to which the driver places inappropriately high trust. This paper distinguishes five types of target objects, computer (C), software (S), hard-ware (H), environment (E), and liveware (L) according to the C-SHEL model [19] describing human interactions with other humans, technology, and the environment; see, Figure 1.

Example 4 (overtrust in computer). The adaptive cruise control system (ACC) performs the longitudinal control on behalf of a driver. Suppose the driver thought that "a car just ahead of me on the next lane may be cutting in. The ACC must have already noticed the car and will adjust the control when appropriate." This is overtrust in the ACC (computer), when the car on the next lane is outside the range of the ACC and the driver does not notice that.

Example 5 (overtrust in software). Imagine a case in which a driver thought that "today is the first day for me to use a brand new system. Oh, I forgot to read the manual. There should be no problem even if I pressed the buttons in a wrong sequence. Fool-proof or tamper-proof functions must be implemented in the software."

Example 6 (overtrust in hardware). Assume that a driver thought that "strictly speaking, this is the time for me to bring the car to a periodic inspection. However, I am quite busy right now and I have never experienced hardware troubles in the car. Why do I have to bring my car periodically for an inspection? My car will not fail."

Example 7 (overtrust in environment). Suppose a man is driving his car thinking that "this road is simply straight. Moreover, there is usually little traffic. It is very relaxing to drive on this simple and somewhat boring road." In reality, environment may alter with time.

Example 8 (overtrust in liveware). Suppose a driver is approaching to an intersection with a blind corner and that an ADAS sets off an alert telling the driver that "a car is approaching to the intersection from the right on the crossing road." The driver cannot see the car himself, because the car is just behind the blind corner. The ADAS generated the alert based on the information obtained via vehicle-to-vehicle or vehicle-to-infrastructure communication technology. Suppose the driver thinks that "I do not see any car. If there is a car, the car will surely yield the right of way, because it is me that is on the priority road," which is overtrust in the driver (liveware) of the other car at the blind corner.

Example 9 (overtrust in liveware). Imagine a car equipped with an electronic stability control system (ESC) that improves stability by applying the brakes to individual wheels when skids or loss of steering control was detected. Suppose the ESC worked at a sharp curve on a slippery road. If the human interface was not properly designed to let the driver know that the ESC was activated, the driver might feel inappropriate confidence on his driving skill, failing to recognize that it was the ESC that assured the stability of the car at the curve. This is a case of overtrust in the driver himself/herself.

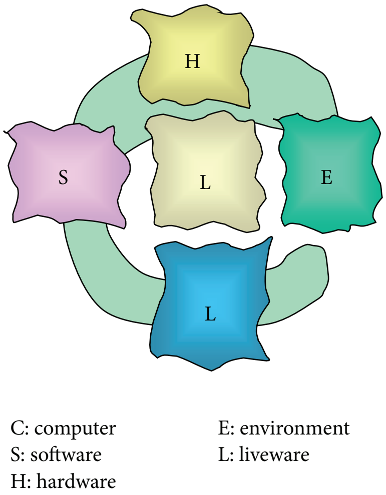


Figure 1: C-SHEL model.

2.3. Chances of Observation. The third axis (1-3) distinguishes two classes for ADAS: (a) ADAS for use in normal driving and (b) ADAS for use in emergency. A most prominent characteristic that distinguishes the two classes is the chances to observe ADAS functioning.

Example 10. ADAS for use in normal driving (e.g., ACC) usually aims to reduce the driver workload and works continuously for certain period of time. Since such an ADAS is used daily, the driver observes the system's "intelligent" behaviors repeatedly, which gives the driver a number of opportunities for constructing a mental model of the ADAS.

Example 11. ADAS for use in emergency (e.g., AEBS described in Section 1) usually aims to prevent a catastrophic event from occurring and thus to attain the driver safety. Since such an ADAS is activated only in cases of emergency, it would be very rare for an ordinary driver to see the ADAS works. That suggests that the driver may not be able to accumulate chances sufficient enough for constructing a concrete mental model of the ADAS.

3. Overreliance

Overreliance on an ADAS is a psychological state in which the human reliance on the ADAS is inappropriately high. More precisely, overreliance is an incorrect action selection decision based on an incorrect situation diagnostic decision regarding the ADAS (i.e., the overtrust in it). Here we introduce three axes for describing types of overreliance, that is, (2-1) type of action selected, (2-2) benefits expected, and (2-3) time allowance for human intervention.

3.1. Type of Action Selected. For the action selection, (2-1), this paper distinguishes the following two types of decisions: commission-like action selection decision and omission-like action selection decision. The former is a selection and implementation of an action that is not suitable to a given situation. Risk compensating behavior [20, 21] could be categorized as a commission-like action. The latter is a failure to select or implement an action that is needed in a given situation.

Example 12 (commission-like action). Suppose a man is driving a car equipped with an ESC at high speeds, which is overreliance on the ESC if it was a clear but extremely cold winter morning and it had rained before dawn. It would be inappropriate to drive a car at high speeds in such an adverse weather condition although the car is equipped with the ESC.

Example 13 (omission-like action). Suppose a man is driving a car by using an ACC and a lane keeping assistance system (LKA). LKA is an automatic system that recognizes the lane and provides the driver with assisting steering torque to keep the car around the center of the lane. Suppose the driver decided to let the LKA take care of the lateral control completely for a while, so that he could consult the navigation system to know how to access his destination. If the LKA was of the type that ceases to control the steering when it determines, through monitoring the driver behavior, that the driver has not been active in steering, the driver's decision to trade the full authority to the LKA is overreliance on the LKA. A case may happen that nobody controls the car, if the human interface did not tell the driver clearly that the LKA returned the authority and responsibility of steering back to the driver based on its decision that the driver had been inactive in steering for a certain period of time.

3.2. Benefits Expected. The second axis is to describe whether the driver can produce some benefits by relying on the assistance system.

Example 14. Suppose the driver assigns the ACC all the tasks for longitudinal control of the vehicle. That may enable the driver to find time to relax muscles and extend legs after stressful maneuvering, or to allocate cognitive resources to find a right way to the destination in a complicated traffic conditions. In this way, relying on the assistance system sometimes brings extra benefit to the driver, when the system is for use in normal driving.

Example 15. AEBS is activated only in emergency, and the time duration for the AEBS to fulfill its function is short, say several seconds. It is thus not feasible for the driver to allocate the time and resources, saved by relying on the AEBS, to something else to produce extra benefit within the several seconds. A similar argument may apply to other assistance systems designed for emergency. If a driver relies on AEBS in normal driving in a sense that the driver lets the AEBS brake when necessary, it would be beneficial for the driver to be able to decrease his/her vigilance and to be relaxed (benefit from an omission-like action). On the other hand, the driver could increase the vehicle speed and reduce time headway and let the AEBS take care of braking. This is another benefit obtained from a commission-like action.

3.3. Time Allowance for Human Intervention. The third axis is to describe whether the driver can intervene into the assistance system's control when the driver determined that the system performance differs from what he or she expected. Note here that this axis could be used to judge whether a driver's reliance is excessive or not. That is, if the time allowance is large enough, the driver has to intervene into control when necessary. However, this axis is not for explaining causes of overreliance.

Example 16. In case of ACC, it may not be hard for the driver to intervene to override the ACC when its performance was not satisfactory. In fact, a driver has to intervene into control when the deceleration of the lead vehicle is larger than what the ACC can manage. If the driver does not apply the brake him/herself, the driver's reliance is regarded as excessive.

Example 17. In case of AEBS, it might be unrealistic to assume that the driver can intervene into control by the AEBS when he or she decided that the AEBS performance was not satisfactory, because the whole process of monitoring and evaluation of AEBS performance as well as decision and implementation of intervention must be done within a few seconds. Thus, driver's failure to override the AEBS when the system is not successful to avoid a crash does not directly mean that the driver's reliance is excessive, if the driver maintained long enough time headway and paid enough attention to the lead vehicle. If, on the other hand, the driver has maintained short time headway and let the AEBS apply the brake, it should be regarded as overreliance.

4. Possibilities of Overtrust and Overreliance

Let us discuss overtrust in and overreliance on ADAS by integrating viewpoints given in Sections 2 and 3.

4.1. Communication-Based Information Provision. Suppose an ADAS has a communication-based function to set off an alert on a car that the driver may not be able to see. There are some objects in which the driver may place overtrust. Example 8 has described one of such cases, where the driver of some other car (a *liveware* in the target-object axis in Section 2.2) needs to be taken into account from a viewpoint of *performance* in the dimension-of-trust axis in Section 2.1.

Consider a case in which a driver is approaching to an intersection that has blind corners but has no traffic lights. The communication-based infrastructure was installed a year ago. The infrastructure can detect cars travelling on the roads crossing each other, and it sends a signal to an onboard ADAS of a car, so that the ADAS can set off an alert to let the driver know an approach or existence of some car(s) on a crossing road. Suppose the driver drives the road daily (i.e., chance-of-observation axis) and has been satisfied with the performance (i.e., dimension-of-trust axis) of the communication-based alert. The driver now thinks that "I am sure that no car is coming toward me when no alert is given. Why not cross the intersection without deceleration?" In this case, the driver is overlooking the possibility of hardware failure of the infrastructure (i.e., target-object axis). His situation diagnostic decision that "no car must be approaching toward me because no alert is there" is inappropriate (i.e., overtrust). When the communication-based infrastructure was out of service, no alert can be given to the driver. Thus, the action selection decision to "cross the intersection without deceleration" is overreliance on the function of the communication-based alert, when the driver abandons the responsibility to be vigilant.

4.2. Adaptive Cruise Control System. Conventional adaptive cruise control (ACC) systems are not able to control headway in reference to slow or stopped vehicles [22, 23]. According to an interview to owners of vehicles equipped with an ACC system, some part of the owners did not understand this inability [24]. Itoh [18] conducted a driving simulator experiment to observe overtrust in and its resulting overreliance on the ACC. Participants were requested to drive a car by using an ACC that can control the host vehicle to a complete stop when the lead vehicle decelerates and stops. However, the ACC does not recognize stationary body (such as, cars standing still). Participants experienced 69 drives with ACC during the period of four days. At the final trial on the fourth day, participants were given a case in which, after 20 minutes of following the lead vehicle at 100 km/h, the lead vehicle made a lane change and the host vehicle happened to approach to the tail of a traffic jam, where all the vehicles in the jam stood completely still. Participants needed to apply the brake by themselves. One collision and some near collisions into the car at the tail of the jam were observed in the experiment. None of the participants who caused the collision or near collision were drowsy or distracted. Data analyses and investigation of those cases suggested that the participants developed trust in the ACC while experiencing repeatedly the ACC's successful lead vehicle followings to complete stops (i.e., chance-of-observation axis), and that some participants had inappropriate expectations (i.e., dimension-of-trust axes) that the "ACC would control the host vehicle nicely to a vehicle ahead," even for an already standing still vehicle. The participants' failure in applying the brake (i.e., omission-like action) is due to overreliance on the ACC, induced by the overtrust in it.

4.3. Airbag. Strictly speaking, airbags may not be ADAS. However, it is worth mentioning that problems are related to deployment of airbags, because the problems are closely related to the issues of overtrust and/or overreliance.

Since 1990s, passenger cars have been widely equipped with airbags. However, related to the use of airbags, many troubles occurred in Japan, especially at the early stage of the spread. For example, there were cases in which a driver was killed or seriously injured by the deployment of the airbag when the vehicle crashed into something (see, e.g., [25]). In those cases, the drivers had not fastened their seat belt when they had the accident. One possible reason for the nonuse of the seat belt was that the drivers regarded the airbag as an alternative to a seat belt. However, such understanding is inappropriate. An airbag is a "supplemental restraint system" (SRS) which means that the airbag is supplemental to the seat belt. An airbag itself is not enough to support the driver. Regarding an airbag as an alternative to the seat belt can be said as overtrust in an airbag system in terms of "purpose" dimension of trust. Note here that such overtrust could emerge even if the driver does not have chances to observe cases when an airbag works at all beforehand. The process of emerging such excessive trust in an airbag system is different from the one for ACC systems. Driver's reliance on an airbag without being belted is overreliance. It is an omission-like action in a sense that the driver omits fastening his/her seatbelt. For drivers, being unbelted may be worth doing because it could be relaxing for the drivers. This type of overreliance can be detected by monitoring the state of the seatbelt.

On the other hand, some drivers complained about the nonactivation of an airbag when their vehicle crashed into something. For most cases, the reason for the non-activation was not a malfunction of the airbag, but the situation was beyond the system's operative conditions. For example, the airbag at the driving seat may not deploy in a case of offset crashes. It can be said that the complaint is due to driver's overtrust in process in the systems if the driver thinks just "I do not know how the airbag system works, but it will deploy whenever a crash occurs." This is another type of overtrust in an airbag system. Note here that the driver may fasten his/her seatbelt even if the driver has such overtrust in an airbag system. This type of overtrust may not be found until a crash occurs.

The above two examples suggest that it is necessary to identify what type of overtrust/overreliance which is under consideration.

4.4. Advanced Emergency Brake System. Conventional AEBS did not aim to prevent a catastrophe from occurring but to mitigate collision damages. Troubles due to drivers' overreliance on AEBS have not been reported from field operations in the real world. On the other hand, development of technologies increased the possibility of AEBS for collision avoidance. Thus, it has been an important question to be addressed: Do drivers place too much trust in and overreliance on AEBS for collision avoidance? The answer can be given by investigating the possibility of the overtrust and the overreliance with the theoretical framework proposed in this paper.

Since the system is activated only in cases of emergency, it would be very rare for an ordinary driver to see how the system works (i.e., chance-of-observation axis). It is thus highly possible that the driver will not be able to construct a precise mental model of the AEBS through the use of it. This suggests that it may be hard for the driver to engender a sense of trust in the system, especially in terms of "performance" (i.e., dimension-of-trust axis). What happens then? No possibility for the driver to place overtrust in the AEBS? The answer may be negative. It is known that people may place inappropriate trust (i.e., overtrust), especially in terms of process and/or purpose dimensions, without having any concrete experience or evidence proving that the object is trustworthy; see, for example, [17]. Our experience with drivers' overtrust in and overreliance on airbag systems also supports the concern with overtrust in AEBS.

Suppose the driver places overtrust in the system. Does that mean that the driver relies on the system too much (i.e., overreliance)? In one sense, the answer may be positive. Itoh et al. [26] conducted a driving simulator experiment and found that drivers shortened the time headway while they followed a lead vehicle when an AEBS for collision avoidance was available. This is an example of commission-like actions. However, such overreliance was partly due to repetitive experience of the AEBS in the experiment. In reality, it is rare for a driver to observe cases where the automatic collision avoidance brake is activated.

In addition, the drivers were not distracted at all in the experiment even when the AEBS was available, and the drivers' reaction against the rapid deceleration of the lead vehicle was not delayed in the experiment of Itoh et al. [26]. That is, the drivers may not rely on the AEBS excessively in a sense that they allocate their resources to something else at the risk of their life (i.e., benefit-expected axis).

In case of an ADAS designed for use in normal driving situations, even if the system's behavior was not what the driver expected, there would be enough time for the driver to override the system to cope with the circumstances himself or herself. However, in case of an ADAS for emergency use, even if the driver noticed that the system's behavior was not what he or she expected, no time may be left for him or her to correct it (i.e., time-allowance-for-human-intervention axis).

The above discussion suggests that AEBS for collision avoidance could be free from drivers' overreliance if the system is designed appropriately.

In Japan, a national advanced safety vehicle (ASV) project discussed this issue. One of the authors of this paper was the leader of the task force in the ASV project to investigate design requirements for AEBS with collision avoidance functionality. As a conclusion, the ASV task force approved that the AEBS may be developed as a collision "avoidance" system, instead of a collision damage mitigation system. Such collision avoidance AEBS may not interfere with the driver's own actions (by letting it apply the automatic brakes at the latest time possible), but still it can avoid a collision against a forward obstacle effectively. Human factors viewpoints played major roles in determining the design requirements on the AEBS timing to initiate an automatic emergency braking and its deceleration rate. In fact, they were determined through the analyses of drivers' braking behaviors in normal and critical traffic conditions. Moreover, a couple of conventional requirements for the AEBS were abolished from human factors viewpoints (e.g., to reduce mode confusion or automation surprises). Based on the conclusion of the ASV task force, the Ministry of Land, Infrastructure, and Transport has been revising the design guidelines for the AEBS.

5. Discussions

This paper has proposed a theoretical framework to discuss the driver's overtrust in and overreliance on ADAS in a precise manner. Overtrust and overreliance are distinguished rigorously, and their characteristics are illustrated by introducing some viewpoints (or aspects and axes). It has been shown that our theoretical frame enables precise description, classification, rigorous analysis, and evaluation of the driver's overtrust in and overreliance on ADAS. Since the framework distinguishes the target object of the driver's overtrust, it can be used to derive a countermeasure for reducing the possibilities of the driver's overtrust. In other words, a systematic investigation can be made possible to determine whether overtrust in question may be alleviated by improving human-machine interface, or by preparing a better operation manual, or by providing the drivers with opportunities to acquire knowledge and/or to improve skills, or by some other means.

It would be apparent that alleviation or prevention of overtrust in or overreliance on the ADAS and its effects on degradation of safety of the car-driver system are closely linked to the issue of authority and responsibility. It is sometimes useful to provide the driver with multilayered assist functions [1]. In the first layer, a driver's situation recognition and understanding are enhanced for proper situation diagnostic decisions and associated action selection decisions. In the second layer, the ADAS monitors the driver's behaviours and traffic conditions to evaluate whether his or her intent and behaviours match the traffic conditions. When the ADAS detects a deviation from normality (for instance, by detecting behaviours or postures that suggest the driver's overtrust or its resulting overreliance), it gives the driver an alert to make him or her return to normality. In the third layer, the ADAS provides the driver with automatic safety control functions, if the deviation from normality still continues to be observed or if little time is left for the driver to cope with the traffic conditions. The situation-adaptive ADAS adjusts its assist functions dynamically, so that they may fit to the human's intent, psychological/physiological conditions, and the traffic conditions. The adjustment of assist functions is made in a *machine-initiated* manner [27–29] by inferring intent and conditions of the human through monitoring his or her behaviours. It is proven mathematically in [29] that a machine-initiated trading of authority based on the machine's interpretation of the situation and the human's behaviour is indispensable for assuring safety of the car-driver system, although the machine-initiated policy is not human-centered in the sense of [4].

The driver's control action may be classified into three categories: (1) an action that needs to be done in a given situation, (2) an action that is allowable in the situation and thus it may either be done or undone, and (3) an action that is inappropriate and thus must not be done in the situation. Assuming sensing technology for the computer (ADAS), two states may be distinguished for each control action: (a) "detected," in which the computer judges that the driver is performing the control action, and (b) "undetected," in which the control action is not detected by the computer (Figure 2).

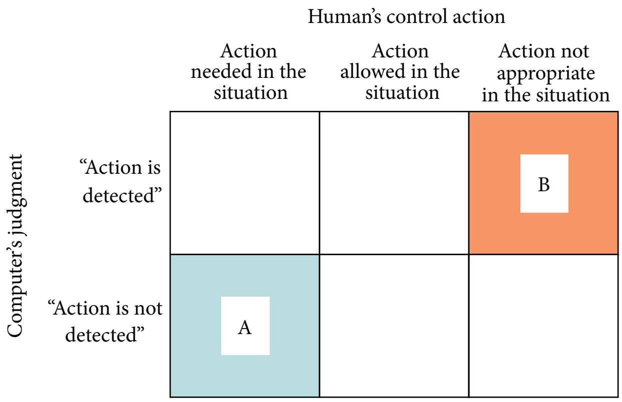


Figure 2: Control action in a given situation.

Case A represents a circumstance with the driver's omission-like action selection, while case B depicts a circumstance with the driver's commission-like action selection and implementation. These mismatches between the driver's action selection decision and the given situation can occur when the driver may place overreliance on the ADAS, as has been discussed already. Then the question becomes, "what is a sensible and effective countermeasure for the ADAS in such circumstances? Is it enough for the ADAS to set off an alert to let the driver resolve the mismatch himself or herself? Or, is it better for the ADAS to initiate an automatic control action to cope with the situation?" Inagaki and his colleagues have shown that the authority may be given to the ADAS, so that (i) it can take an automatic safety control action that the driver failed to perform, or (ii) it can take a protective action (*soft protection* or *hard protection*) that tries to prevent the driver's inappropriate action causing an accident or an incident [30–32].

References

[1] T. Inagaki, "Smart collaboration between humans and machines based on mutual understanding," *Annual Reviews in Control*, vol. 32, no. 2, pp. 253–261, 2008.

[2] E. Hollnagel and A. Bye, "Principles for modelling function allocation," *International Journal of Human Computer Studies*, vol. 52, no. 2, pp. 253–265, 2000.

[3] G. Klein, "A recognition-primed decision (RPD) model of rapid decision making," in *Decision Making in Action*, G. Klein, J. Orasanu, and R. Calderwood, Eds., pp. 138–147, Ablex, 1993.

[4] C. E. Billings, *Aviation Automation—The Search for a Human-Centered Approach*, LEA, 1997.

[5] T. Inagaki, "Design of human-machine interactions in light of domain-dependence of human-centered automation," *Cognition, Technology and Work*, vol. 8, no. 3, pp. 161–167, 2006.

[6] M. R. Endsley and E. O. Kiris, "The out-of-the-loop performance problem and level of control in automation," *Human Factors*, vol. 37, no. 2, pp. 381–394, 1995.

[7] N. B. Sarter and D. D. Woods, "How in the world did we ever get into that mode? Mode error and awareness in supervisory control," *Human Factors*, vol. 37, no. 1, pp. 5–19, 1995.

[8] R. Parasuraman and V. Riley, "Humans and automation: use, misuse, disuse, abuse," *Human Factors*, vol. 39, no. 2, pp. 230–253, 1997.

[9] N. B. Sarter, D. D. Woods, and C. E. Billings, "Automation surprises," in *Handbook of Human Factors and Ergonomics*, G. Salvendy, Ed., pp. 1926–1943, John Wiley & Sons, 2nd edition, 1997.

[10] T. Inagaki and J. Stahre, "Human supervision and control in engineering and music: similarities, dissimilarities and their implications," *Proceedings of the IEEE*, vol. 92, no. 4, pp. 589–600, 2004.

[11] R. Parasuraman, R. Molloy, and I. L. Singh, "Performance consequences of automation-induced 'complacency'," *International Journal of Aviation Psychology*, vol. 3, no. 1, pp. 1–23, 1993.

[12] K. L. Mosier, L. J. Skitka, S. Heers, and M. Burdick, "Automation bias: decision making and performance in high-tech cockpits," *International Journal of Aviation Psychology*, vol. 8, no. 1, pp. 47–63, 1998.

[13] J. Meyer, "Effects of warning validity and proximity on responses to warnings," *Human Factors*, vol. 43, no. 4, pp. 563–572, 2001.

[14] T. B. Sheridan and R. Parasuraman, "Human-automation interaction," in *Reviews of Human Factors and Ergonomics*, R. S. Nickerson, Ed., vol. 1, pp. 89–129, HFES, 2005.

[15] *Complacency*, Merriam-Webster Online Dictionary, 2010.

[16] N. Moray and T. Inagaki, "Attention and complacency," *Theoretical Issues in Ergonomics Science*, vol. 1, no. 4, pp. 354–365, 2001.

[17] J. Lee and N. Moray, "Trust, control strategies and allocation of function in human-machine systems," *Ergonomics*, vol. 35, no. 10, pp. 1243–1270, 1992.

[18] M. Itoh, "Toward overtrust-free advanced driver assistance systems," *Cognition, Technology, and Work*, vol. 14, no. 1, pp. 51–60, 2012.

[19] T. Inagaki, "New challenges on vehicle automation: human trust in and reliance on adaptive cruise control systems," in *Proceedings of the IEA (CD-ROM)*, p. 4, 2003.

[20] G. J. S. Wilde, *Target Risk*, PDE Publications, Toronto, Canada, 1994.

[21] OECD Road Transport Research, *Behavioral Adaptations to Changes in the Road Transportation System*, OECD, Paris, France, 1990.

[22] M. Weinberger, H. Winner, and H. Bubb, "Adaptive cruise control field operational test—the learning phase," *JSAE Review*, vol. 22, pp. 487–494, 2001.

[23] B. D. Seppelt and J. D. Lee, "Making adaptive cruise control (ACC) limits visible," *International Journal of Human Computer Studies*, vol. 65, no. 3, pp. 192–205, 2007.

[24] D. A. Dickie and L. N. Boyle, "Drivers' understanding of adaptive cruise control limitations," in *Proceedings of the 53rd Human Factors and Ergonomics Society Annual Meeting (HFES '09)*, pp. 1806–1810, October 2009.

[25] K. Tanno, M. Kohno, K. Ono et al., "Fatal cardiovascular injuries to the unbelted occupant associated with airbag deployment: two case-reports," *Legal Medicine*, vol. 2, no. 4, pp. 227–231, 2000.

[26] M. Itoh, Y. Fujiwara, and T. Inagaki, "Driver behavioural changes through interactions with an automatic brake systems," *Transactions of the Society of Instrument and Control Engineers*, vol. 47, no. 11, pp. 512–519, 2011 (Japanese).

[27] T. Inagaki, "Adaptive automation: sharing and trading of control," in *Handbook of Cognitive Task Design*, E. Hollnagel, Ed., pp. 147–169, LEA, 2003.

[28] M. W. Scerbo, "Theoretical perspectives on adaptive automation," in *Automation and Human Performance*, R. Parasuraman and M. Mouloua, Eds., pp. 37–63, LEA, 1996.

[29] T. Inagaki and T. B. Sheridan, "Authority and responsibility in human-machine systems: probability theoretic validation of machine-initiated trading of authority," *Cognition, Technology & Work*, vol. 14, no. 1, pp. 29–37, 2012.

[30] T. Inagaki, M. Itoh, and Y. Nagai, "Efficacy and acceptance of driver support under possible mismatches between driver's intent and traffic conditions," in *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society (HFES '06)*, pp. 280–283, October 2006.

[31] T. Inagaki, M. Itoh, and Y. Nagai, "Driver support functions under resource-limited situations," in *Proceedings of the 51st Annual Meeting of the Human Factors and Ergonomics Society (HFES '07)*, pp. 176–180, October 2007.

[32] T. Inagaki, M. Itoh, and Y. Nagai, "Support by warning or by action: which is appropriate under mismatches between driver intent and traffic conditions?" *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E90-A, no. 11, pp. 2540–2545, 2007.

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2013, Article ID 912860, 9 pages
http://dx.doi.org/10.1155/2013/912860

Research Article

Predicting Driver Behavior Using Field Experiment Data and Driving Simulator Experiment Data: Assessing Impact of Elimination of Stop Regulation at Railway Crossings

Toshihisa Sato,[1] Motoyuki Akamatsu,[1] Toru Shibata,[2] Shingo Matsumoto,[2] Naoki Hatakeyama,[2] and Kazunori Hayama[2]

[1] *Human Technology Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Central 6, 1-1-1 Higashi, Tsukuba 305-8566, Japan*
[2] *Human Science Division, Railway Technical Research Institute, 2-8-38 Hikari-cho, Kokubunji-shi, Tokyo 185-8540, Japan*

Correspondence should be addressed to Toshihisa Sato; toshihisa-sato@aist.go.jp

Received 14 September 2012; Accepted 1 March 2013

Academic Editor: Klaus Bengler

Copyright © 2013 Toshihisa Sato et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We investigated the impact of deregulating the presence of stop signs at railway crossings on car driver behavior. We estimated the probability that a driver would stop inside the crossing, thereby obstructing the tracks, when a lead vehicle suddenly stopped after the crossing and a stop regulation was eliminated. We proposed a new assessment method of the driving behavior as follows: first, collecting driving behavior data in a driving simulator and in a real road environment; then, predicting the probability based on the collected data. In the simulator experiment, we measured the distances between a lead vehicle and the driver's vehicle and the driver's response time to the deceleration of the leading vehicle when entering the railway crossing. We investigated the influence of the presence of two leading vehicles on the driver's vehicle movements. The deceleration data were recorded in the field experiments. Slower driving speed led to a higher probability of stopping inside the railway crossing. The probability was higher when the vehicle in front of the leading vehicle did not slow down than when both the lead vehicle and the vehicle in front of it slowed down. Finally, advantages of our new assessment method were discussed.

1. Introduction

Driving simulators have been used to evaluate driver behaviors and the influences of advanced driver assistance systems on the driving behaviors. The simulators are an essential tool in automotive human factors research. Advantages of using the driving simulator are a safety (no traffic accidents during the driving experiments), an easy collection of the driver behavior data, and a precise reproduction of road traffic environments within and between the drivers. The driving simulator has contributed to collecting driving behavior data under situations with potential risks for a traffic accident (e.g., situations where pedestrians or bicycles suddenly rush out in front of the driver's vehicle). The data collection and an assessment of the driver behavior when avoiding the incidents have led to developing advanced driver assistance systems and warning systems for reducing these accidents. However, the experiments under the risky conditions using the simulators have mainly two disadvantages: (1) a driving simulator cannot fully reproduce the driver behavior on a real road, and (2) once a driver experiences a target situation, he/she becomes cautious about similar situations and the behavior data under the same target situation and/or the other risky situations with different traffic conditions will never be collected.

Recent technologies have been applied to the developments of driving simulators and improved driver's feelings while driving the simulators [1, 2]. However, the driving simulator cannot fully reproduce the driving behavior in an actual road environment due to mechanical restrictions. Particularly, deceleration is one of the difficult aspects in improving the driver's feelings, because it is impossible to reproduce the same amount of deceleration using the simulator as in real-world braking. This is due to the lack of motion in fixed-base simulators and the impossibility of reproducing full longitudinal movement in motion-based simulators. Stopping at driver's anticipating point is a difficult task, although several studies have attempted to clarify the factors influencing braking maneuvers in a driving simulator [3, 4].

Data collection under a situation where drivers might experience traffic accidents is one of the merits of a driving simulator experiment. We can reproduce the same traffic situations any number of times using the simulators; we cannot, however, restore the driver's experiences of facing the risky conditions. Once a driver experiences a traffic accident or its near miss in a simulator experiment, he/she becomes cautious about similar unsafe situations after the event. Therefore, we cannot collect driving behavior data under several kinds of situations leading to traffic accidents from one participating driver, and more and more participants are needed to evaluate the driver behavior under such situations.

We propose a new assessment method of the driver behavior in order to overcome these disadvantages. We evaluate the driver behavior that is estimated based on experimental data, instead of evaluating a real behavior observed in an experiment. In the estimation, we use field experiment data for the driving behavior that a driving simulator cannot fully reproduce, in addition to the driving behavior data which can be collected in the simulator environment. In the driving simulator experiment, the participating driver does not experience a target situation in which a traffic accident occurs, and he/she experiences a situation just before the target situation will happen. We can collect driving behavior data under various kinds of road traffic situations from one participant, which are used for the estimation, because the participant does not receive feedback about his/her exposure to the target situation. The proposed method was applied to an assessment of impact of eliminating stop regulation at railway crossing on the driver's unexpected stopping on the train track.

Article 33 of the Road Traffic Act of Japan stipulates that all vehicles must stop before driving through railway crossings. However, this obligation to always stop before entering crossings is one of the causes of problems such as all-day traffic congestion due to poor traffic flow efficiency. One countermeasure that can be taken to resolve this problem is to amend the law to remove the obligation to stop before entering railway crossings. Removal of the need for an obligatory stop is expected to alleviate problems such as traffic congestion. Although this countermeasure is expected to reduce traffic congestion, we are afraid that the change may increase the likelihood of vehicles becoming stuck in railway crossings (in the field of railway research in Japan, this condition is called "trapped"): drivers may be more likely to enter the crossing without checking whether or not there is enough space on the other side, and they may become trapped if vehicles ahead of them are stopped. For this reason, before the law is changed, it is essential to assess the probability of drivers becoming trapped in crossings. In this research, we investigated how the likelihood of drivers becoming trapped was affected by the scenario, including the speed of the vehicle when entering the crossing and the movements of lead vehicles.

Most people think that the obligatory stop will be replaced by an obligation to drive very slowly. In this experiment, we examined the probability of drivers becoming trapped at a much slower vehicle speed of 10 or 20 km/h (=very slow driving conditions). We then compared this with the probability of becoming trapped at a slow or moderate vehicle speed (30 or 50 km/h, resp.).

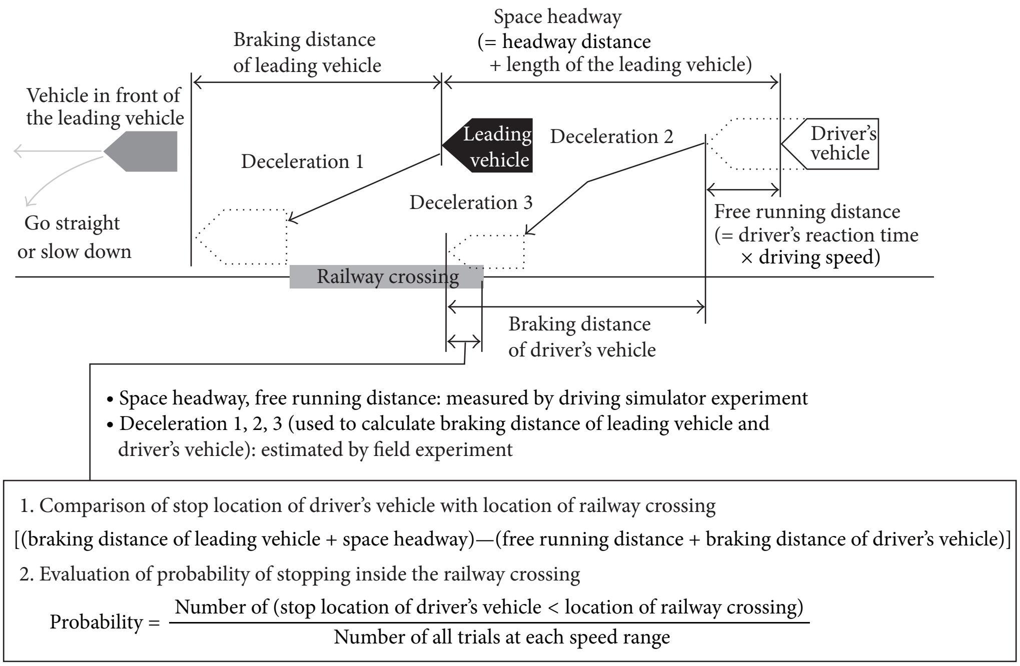
It is considered that a condition where there is no space on the other side of the railway crossing will occur when a driver follows lead vehicles in congested traffic flow. We hypothesized that a driver glances at and recognizes movements of vehicles in front of a leading vehicle and makes a decision on their actions under such car-following conditions. In the experiment, we focused on the presence of two vehicles in front of the driver's vehicle. In this scenario, it is possible that both lead vehicles will decelerate and stop on the other side of the railway crossing. Alternatively, if the road happens to have an intersection immediately after the railway crossing, it is possible that the vehicle in front of the lead vehicle will not decelerate but the lead vehicle will decelerate and/or stop to make a left or right turn. Thus, we investigated the influence of the behavior of the vehicle in front of the lead vehicle on the probability of a driver becoming trapped.

2. Method of Evaluating the Probability of Becoming Trapped

Because drivers in Japan are currently obliged to stop before entering a railway crossing, it is impossible to perform a real-life crossing study in which the stop sign is ignored. In addition, as mentioned in Section 1, the event of becoming trapped will never occur, once a driver experiences the situations where a lead vehicle suddenly decelerates and stops just after the railway crossing during experimental trials.

Therefore, instead, we simulated the stopping position of the driver's vehicle at the time when the lead vehicle happened to decelerate and stop immediately after driving through the railway crossing; we then evaluated the probability of the driver's vehicle becoming trapped by comparing the calculated stop position of the driver's vehicle and the location of the railway crossing. The values of the various parameters required for the simulation were collected from the results of actual road experiments and driving simulator experiments.

Figure 1 presents the method used to evaluate the probability of becoming trapped. The following four parameters were used to calculate the stopping location of the driver's vehicle: the braking distance of the lead vehicle when it decelerated at a rate such that it stopped immediately after driving through the railway crossing (Deceleration 1 in Figure 1); the space headway (=headway distance (distance between the rear of the leading vehicle and the front of the driver's vehicle) + length of the leading vehicle) at the time of the start of the abovementioned deceleration; the free running distance during the driver's brake response time (=the time the driver reacted to the deceleration of the leading vehicle × the driving speed); and the braking distance of the driver's vehicle from the time the vehicle began to decelerate at a specified rate to when it stopped (Deceleration 2 and 3 in Figure 1).



- Space headway, free running distance: measured by driving simulator experiment
- Deceleration 1, 2, 3 (used to calculate braking distance of leading vehicle and driver's vehicle): estimated by field experiment

1. Comparison of stop location of driver's vehicle with location of railway crossing
[(braking distance of leading vehicle + space headway)—(free running distance + braking distance of driver's vehicle)]
2. Evaluation of probability of stopping inside the railway crossing

$$\text{Probability} = \frac{\text{Number of (stop location of driver's vehicle} < \text{location of railway crossing)}}{\text{Number of all trials at each speed range}}$$

Figure 1: Methods used to evaluate the probabilities of the test vehicle becoming trapped inside the railway crossing.

The deceleration rates needed to calculate the braking distances of the driver's vehicle and the lead vehicle were obtained from driving behavior data collected on actual roads. In the case of vehicle speeds of 10 and 20 km/h, we used driving behavior data collected from actual road experiments in Tokyo. In the case of vehicle speeds of 30 and 50 km/h, we collected the driving behavior data on driving routes in Tsukuba. The deceleration rates of the vehicle were extracted from the collected data in which the vehicle slowed down to a stop behind a leading vehicle at a traffic stop light. The deceleration rates while stopping behind the leading vehicle at the traffic stop light are considered to be similar to the deceleration rates obtained while stopping behind a lead vehicle before the railway crossing.

To obtain the braking distance of the driver's vehicle, two deceleration rates were configured for this vehicle: first, it was assumed that the driver would begin to decelerate at a rate that was relative to the deceleration rate of the leading vehicle before entering the railway crossing. Second, before the driver entered the crossing and from the moment the driver recognized that the lead vehicle's driver's intent was to stop, the driver would begin to decelerate at a fairly sharp rate to avoid entering the crossing. Because in almost all of the scenarios from which the actual road driving behavior data were obtained there was a vehicle in front of the lead vehicle, the deceleration rates were used in the simulation with a lead vehicle and another vehicle in front of it.

Both the space headway between the lead vehicle and the driver's vehicle and the driver's free running distance were obtained through driving experiments conducted on the driving simulator. We measured (1) the driver's space headway when the driver's vehicle was following the lead vehicle, at the deceleration point at which the lead vehicle started to decelerate so as to stop immediately after driving through the railway crossing and (2) the driver's brake response time when the lead vehicle was decelerating, with brake lights on, while driving through the crossing. By using this brake response time and the vehicle speed when the driver responded to the brake lights of the lead vehicle, we calculated the driver's free running distance.

3. Measurement of Space Headway and Brake Response Time via the Driving Simulator Experiment

3.1. Test Environment. Figure 2 presents an overview of AIST driving simulator used in the simulator experiment. This driving simulator was configured with a real vehicle cabin, an environment with a 300-degree field of view, a six degree-of-freedom electric motion system, and a sound system with eight spatial placed speakers so that the driver could feel as if they were driving a real vehicle. (For details of the motion system and sound system, refer to [5].) The front-view images displayed on the circular screen were projected by three projectors, but for the central-view area another projector was used to display images with higher resolution. The driver in the vehicle cabin could easily identify the distance to a railway crossing and could recognize the illumination of the lead vehicle's brake lights and the vehicle's deceleration behavior, even if the vehicle was about 100 m away. The driver operated a steering wheel, an acceleration pedal, and a brake pedal. Driving behavior data were recorded at a sampling rate of 60 Hz along with other vehicle data such as speed, position coordinates of the driver's vehicle, and distance to the lead vehicle.

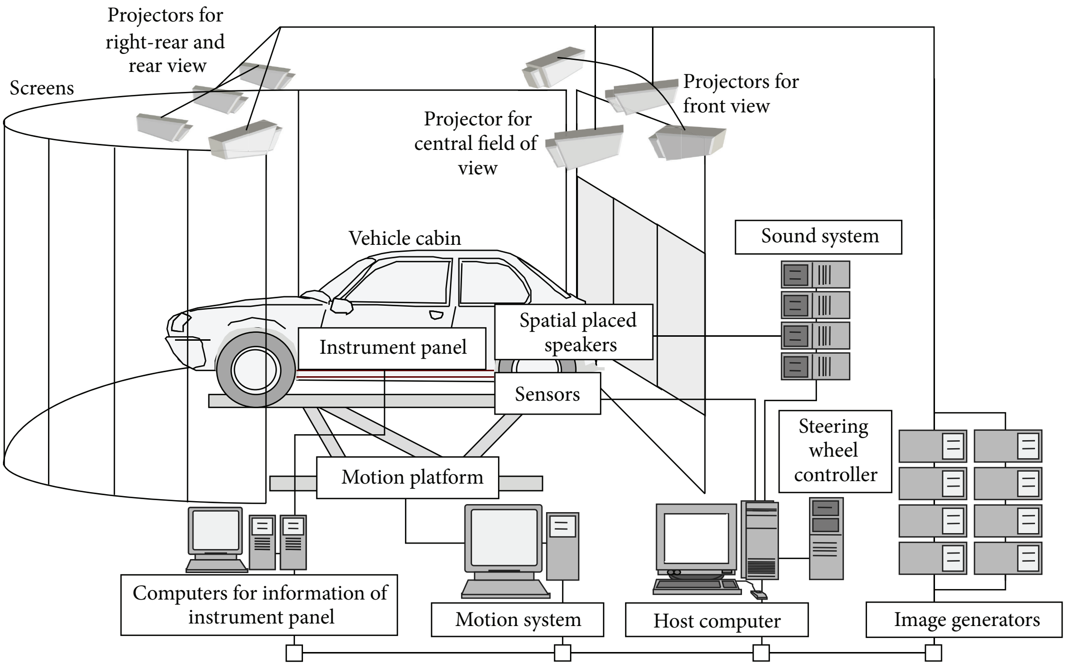


Figure 2: Overview of AIST driving simulator.

Figure 3 presents the driving course and the railway crossing environment used in the experiment. The road had two lanes in each direction, where driving is on the left side of the road. The railway crossing was located about 1300 m from the starting point; after a downhill stretch of 500 m, it went through a railway crossing about 200 m ahead. The railway crossing had one track in each direction, which is standard in Japan. To prevent poor visibility of the railway crossing from affecting the test results, a train was made to pass through the railway crossing. This train could be seen from far away, while the driver was driving downhill, so that the driver could recognize in advance that the crossing was ahead.

3.2. Experiment Participants. A total of 30 drivers (15 males and 15 females) participated in the driving simulator experiment. Their average age was 49.7 years (from 20 to 71), and they had been driving for an average of 21.3 years (from 2 to 50 years). The average annual driving distance was 22,400 km (from 3000 to 50,000 km). Most participants drove nearly every day; participants who had the least driving frequencies still drove 1 or 2 days a week (two participants).

It was explained to participants both verbally and in writing that the purpose of the research was removal of the obligatory stop at railway crossings. They were also told the objective of the driving simulator test, the test content, and the data handling policy. They were told that they were free to withdraw their consent and discontinue participation at any time. After being informed, they were asked for their consent in writing.

3.3. Traffic Conditions. The following three behaviors of the lead vehicle and the vehicle in front of it were prepared: (1) the two vehicles drove through the railway crossing without slowing down (2) both vehicles slowed down (as evidenced by brake lights illumination) before entering the railway crossing and (3) only the lead vehicle slowed down (i.e., the brake light was illuminated) before entering the railway crossing. In scenario (2), the lead vehicle began to slow down approximately 1 s after the vehicle in front of it began to slow. As shown in Figure 3, the lateral positions of the lead vehicle and the vehicle in front of it were shifted so that the test driver could clearly acknowledge the presence of another vehicle in front of the lead vehicle and the illumination of its brakes, because the glass windows of the vehicles included in the simulator image database were not transparent.

To avoid simulator sickness due to sudden stops as well as to avoid giving the feedback to the participants about their behaviors resulting from the lead vehicle's deceleration and stop, after the lead vehicle had slowed down to about half the initial speed, it was accelerated to its original speed and then driven on without stopping.

The driving speeds of the other vehicles were set to very slow (10 and 20 km/h), slow (30 km/h), and moderate (50 km/h). Under each vehicle speed condition, the speed was not always set to be constant but was changed in the range of ±2 km/h.

The total number of test conditions undergone by each test participant was 3 other vehicles' decelerating conditions × 4 vehicle speed conditions = 12 conditions.

3.4. Experiment Procedures. Each test condition was driven once in 1 day. The twelve conditions were driven by the test participants in a random order. The total test period was 3 days, on each test day. Test participants were told that they did not need to stop before entering the railway crossing and that they were to drive in the left lane so as not to pass the lead

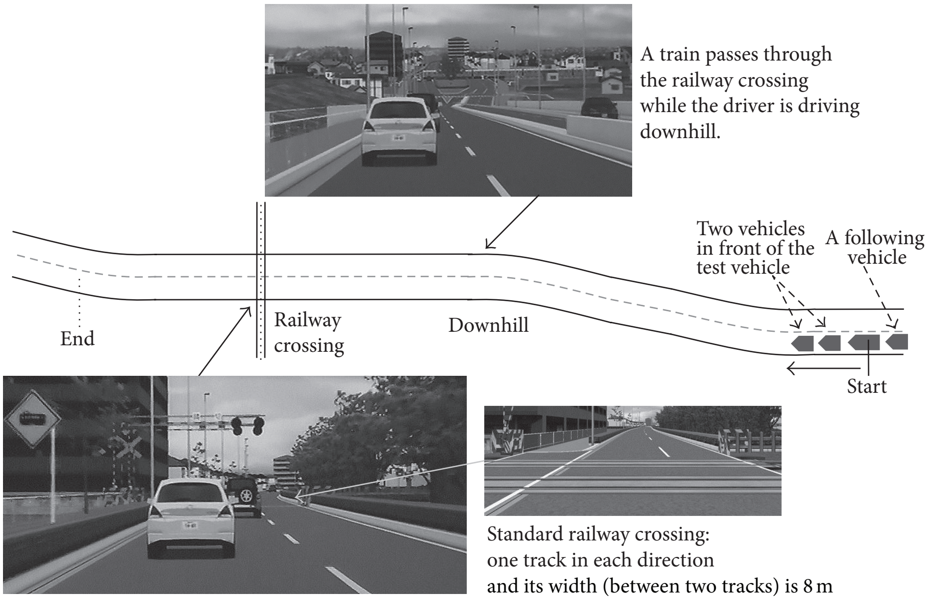


Figure 3: Driving course including railway crossing.

vehicle. They were also instructed to follow the lead vehicle within a reasonable distance and to minimize the headway distance to the lead vehicle until they reached the bottom of the downhill (from then on, they were free to change the headway distance as they wished).

3.5. Measured Items and Results. Data on the space headway between the lead vehicle and the driver's vehicle were measured under only two different traffic conditions where both a lead vehicle and another vehicle in front of it or only the lead vehicle decelerated before entering the railway crossing. We extracted the space headway under each condition at the computed starting point of the lead vehicle's deceleration (as shown in Figure 1).

Two-way analysis of variance (ANOVA) was conducted in which the dependent variable was the space headway and the independent variables were the other vehicle's decelerating conditions (2) and vehicle speed conditions (4). The result showed significant main effects of the other vehicle's decelerating conditions and the vehicle speed conditions (other vehicle's decelerating conditions: $F(3,704) = 413.152$, $P < 0.001$; vehicle speed conditions: $F(1,704) = 4.791$, $P < 0.05$). The post hoc test (Scheffe's method) suggested significant differences between speed conditions. Figure 4 presents the average space headways and their standard deviations. The space headway at 10 km/h condition was the shortest among the four speed conditions. The space headways became longer as the vehicle speed increased. The space headways when both a lead vehicle and another vehicle in front of it slowed down were significantly longer than those when only the lead vehicle slowed down at speed conditions 10 and 20 km/h.

We measured the response time between the first illumination of the lead vehicle's brake lights and the time when the driver first began to brake under conditions in which both the lead vehicle and the one in front of it slowed down and in which only the lead vehicle slowed down. Two-way ANOVA was conducted in which the dependent variable was the brake response time and the independent variables were the other vehicle's decelerating conditions (2) and vehicle speed conditions (4). The result showed a significant main effect for the vehicle speed conditions ($F(3,704) = 3.219$, $P < 0.05$) and a significant interaction between the other vehicle's decelerating conditions and the vehicle speed condition ($F(3,704) = 4.750$, $P < 0.01$). Figure 5 presents the average brake response time and their standard deviations. When the two vehicles slowed down, the driver's brake response time at 50 km/h was earlier than at 30 km/h. The brake response time at 20 km/h was significantly earlier than that at 50 km/h when only the lead vehicle slowed down.

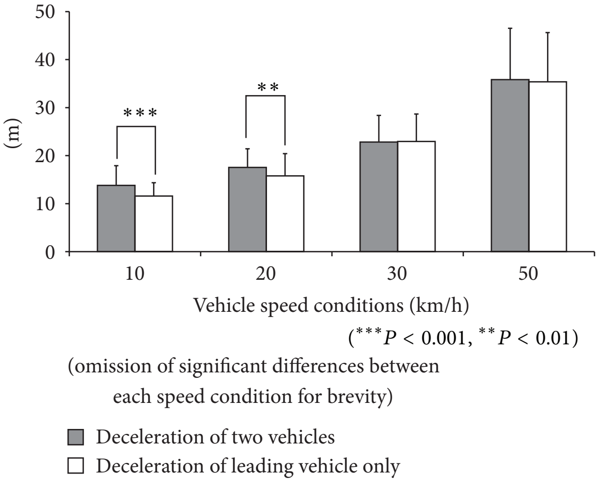


Figure 4: Results of space headway when following two vehicles.

3.6. Space Headway and the Response Time Used in the Simulation. The data of each participating driver was used to compute the probability of the driver becoming trapped. We used the space headway data and brake response data under each condition, because the space headway and brake response time varied according to the behaviors of the lead vehicle and the vehicle in front of it.

4. Collection of Data on Deceleration Rates on Actual Roads

4.1. Collection Method. AIST instrumented vehicle was used to collect deceleration data on actual roads. The instrumented vehicle was equipped with various sensors that detect vehicle velocity by speed pulse signals and deceleration by G-sensor (for details of the instrumented vehicle, see [5]). A driver drove along a predetermined route individually, and data on the driver's usual driving behavior were measured and recorded.

To gather data for vehicle speeds of 10 and 20 km/h, field experiments were conducted in metropolitan Tokyo. At an intersection on an urban road with a speed limit of 30 km/h, we extracted the deceleration rate of the driver's vehicle when the driver followed a lead vehicle and slowed down and then stopped at a stop light. Eight drivers (four males and four females) participated in the Tokyo experiment (average age 42.4 years). The actual average speed at the time of the start of deceleration was 23.2 km/h (range, 6 to 29 km/h).

To gather data for vehicle speeds of 30 and 50 km/h, we carried out field experiments in Tsukuba. The experimental procedures were the same as the experiments in Tokyo. We extracted the deceleration rates of the driver's vehicle at the time when it slowed down behind a lead vehicle and stopped at a traffic stop light. An intersection on a rural road (one lane in each direction) with a speed limit of 40 km/h was selected. Eight drivers with an average age of 37.6 years (four males and four females) participated in the Tsukuba experiments. The actual average speed at the time of the start of deceleration was 35.9 km/h (range, 23 to 48 km/h).

4.2. Results Collected. We examined the distributions of the deceleration rates collected from the experiments in Tokyo (very slow velocity, Figure 6(a)) and the deceleration rates collected from the field experiments in Tsukuba (slow and moderate speed, Figure 6(b)). Both graphs show a cumulative distribution of deceleration rates starting from strong deceleration to no deceleration. All eight drivers' acceleration and deceleration rates were collected via the vehicle's accelerometer at 30 Hz sampling rate, beginning at the time when the driver first began to brake and ending when they reached a vehicle speed of 0 km/h. Because only deceleration was relevant, the negative measurement values were selected. All deceleration rate data were included so as to show the possible distribution range of deceleration rates.

In Figure 6(a), approximately 10% of deceleration rates in the range from strong to zero deceleration were 2 m/s^{2}or stronger. The 25th percentile value of the deceleration rate was about 1.5 m/s^2; the remaining 75% or so of rates were in the 1.5 to 0 m/s^{2}deceleration range. In Figure 6(b), approximately 10% of all deceleration rates were 0.9 m/s^{2}or stronger, and the 25th percentile deceleration value was about 0.6 m/s^2. The results showed that urban drivers whose driving speeds were slower when they first started to decelerate had smaller deceleration rates compared to the rural drivers.

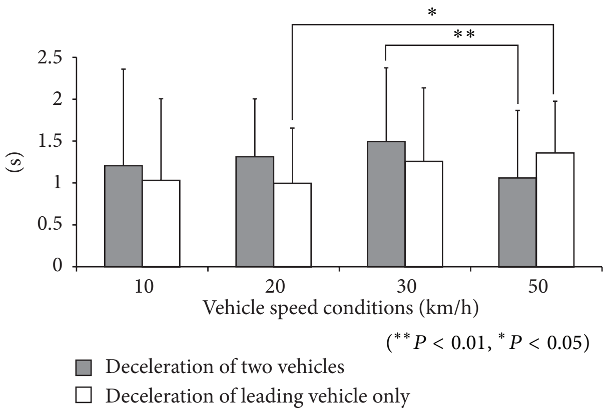


FIGURE 5: Results of driver's brake response time when following two vehicles.

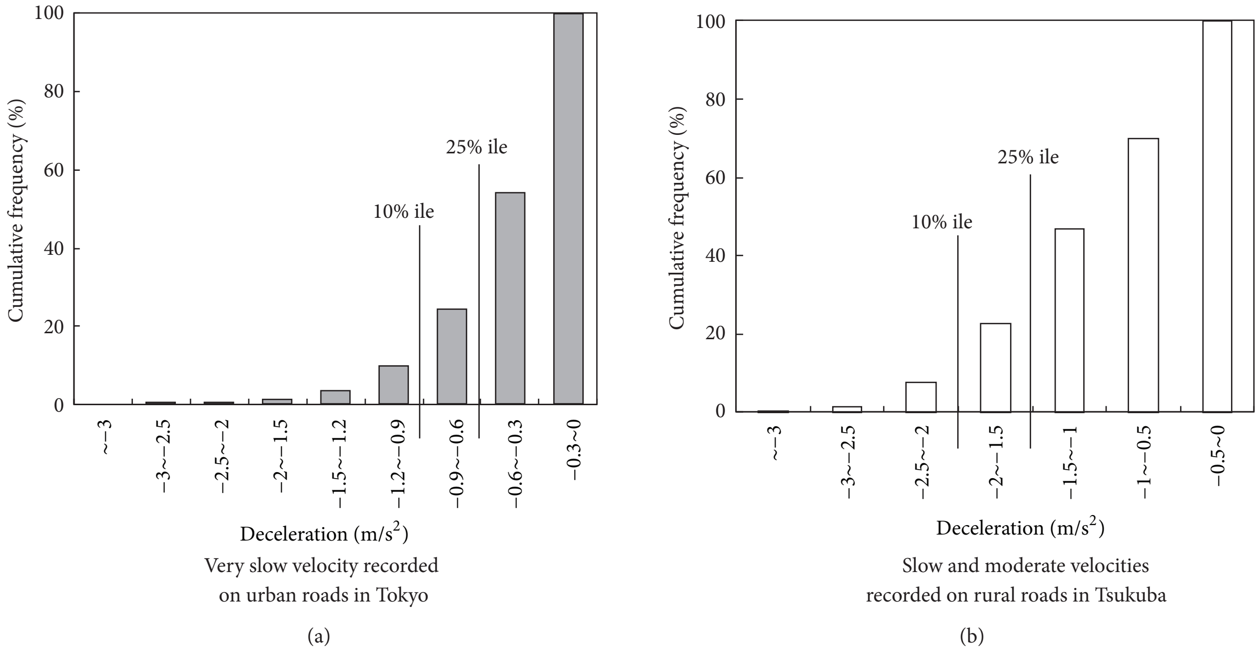
4.3. Extraction of Deceleration Rates for the Simulation. We chose the rates at the 10th and 25th percentile values of the cumulative distribution of deceleration rates collected on actual roads, in order to determine the deceleration rates of the lead vehicle and the driver's vehicle used in the simulation. These rates are presented as follows.

(i) Deceleration rate of the lead vehicle. We chose approximately the 25th percentile value of the deceleration rate (from strong to no deceleration) on actual roads. Under very slow driving conditions, the deceleration rate was set at 0.6 m/s^2, and under slow and moderate conditions, it was set at 1.5 m/s^2.

(ii) Normal deceleration rate of the driver's vehicle. We chose an average deceleration rate between approximately the 10th and 25th percentile values (from strong to no deceleration) on actual roads. Under very slow driving conditions, the deceleration rate was set at 0.73 m/s^2, and under slow and moderately slow conditions, it was set at 1.72 m/s^2.

(iii) Strong deceleration rate of the driver's vehicle. We chose an average deceleration rate in the range between 0 and approximately the 10th percentile value (from strong to no deceleration) on actual roads. Under very slow driving conditions, the deceleration rate was set at 1.21 m/s^2, and under slow and moderate conditions, it was set at 2.33 m/s^2.

The ratio of the time of the normal deceleration rate to the time of the strong deceleration rate was set at 9 : 1, because the strong deceleration was about 10% ile value collected on actual roads from when the driver first began to brake to when they reached a vehicle speed of 0 km/h.



Very slow velocity recorded on urban roads in Tokyo

(a)

Slow and moderate velocities recorded on rural roads in Tsukuba

(b)

Figure 6: Cumulative frequency distribution of deceleration in field experiments.

5. Results of Evaluation of the Probability of Entrapment

We used deceleration rate data collected on actual roads, the space headway measured via the driving simulator experiment, and the data on brake response time, in order to obtain the stop position of the driver's vehicle under each traffic condition and each speed condition. The probabilities of becoming trapped were computed on the basis of a comparison of these positions with the location of the railway crossing. Figure 7 presents the results of the probability of entrapment when both the lead vehicle and the one in front of it slowed down and when only the lead vehicle slowed down. Test vehicles traveling at the slowest speed (10 km/h) had the highest probability of becoming trapped. The probabilities were lower for vehicle speeds of 20 and 30 km/h; at the fastest speed (50 km/h) drivers had the lowest probability of becoming trapped. These tendencies were the same between the two traffic conditions. At every vehicle speed, there was a greater probability of the test vehicle being trapped when only the lead vehicle slowed down than when both the leading vehicle and the vehicle in front of it slowed down. These results were especially found at 10 and 20 km/h.

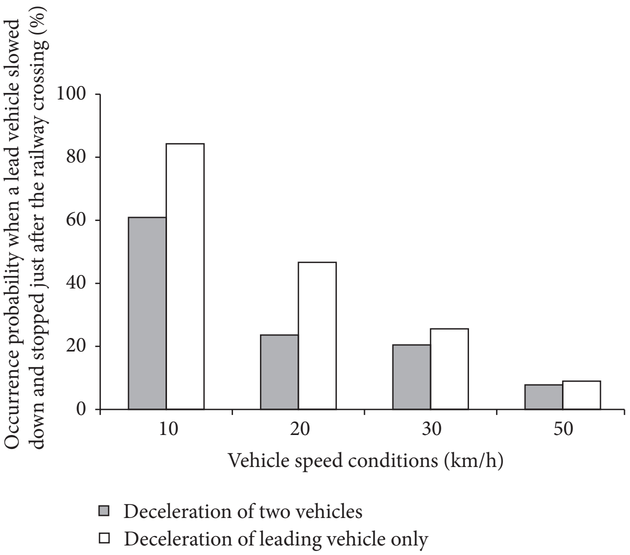


Figure 7: Probability of stopping inside the railway crossing when following two vehicles.

6. Discussion

6.1. Influences of Driving Speeds on the Probability of Entrapment. At a driver's vehicle speed condition of 10 km/h, there was a much higher probability of becoming trapped than at other speeds under both traffic conditions. The space headway available at 10 km/h was the smallest among those for all vehicle speeds.

A speed of 10 km/h can be considered almost equivalent to creeping speed in automatic vehicles. At this speed, drivers are likely to consider that they can stop in no time, and it is therefore conceivable that drivers chose close headway distances to the lead vehicle while approaching the railway crossing. The driver's braking response time tended to become shorter in proportion to the reduction of the space headway, and the driver began to brake significantly more quickly at 20 km/h than at 50 km/h when only a lead vehicle decelerated before entering the crossing. However, the brake response time at 10 km/h was almost the same as at 20 km/h in the case of the deceleration of only the leading vehicle, while the space headway at 10 km/h was significantly shorter compared to the other driving speed conditions. At the slowest speed, drivers chose to simply observe the lead vehicle behavior when only the lead vehicle began to decelerate until they felt it necessary to slow down. In vehicles driving very slowly, the probability of becoming trapped can be the highest because of the combined effect of the small space headway before entry to the railway crossing and the delayed driver's response.

6.2. Influences of Traffic Conditions on the Probability of Entrapment. The probability of the test vehicle becoming trapped was higher when only the lead vehicle slowed down than when both of these vehicles slowed down. This tendency was especially strong at 10 and 20 km/h. This was likely caused by the fact that the space headway at 10 and 20 km/h was significantly smaller when only the lead vehicle slowed down than when the lead vehicle and the one in front of it slowed down.

One reason for this larger space headway when there were two vehicles in front decelerating than when there was only one is because in the former case the driver, seeing the brake lights of the vehicle in front of the lead vehicle go on, would have released the accelerator pedal and already have begun to decelerate before the lead vehicle's deceleration. This feature can be found in the driver's brake response time at 50 km/h when both two vehicles decelerated, which was earlier than when only a leading vehicle decelerated.

The driver pays attention to the behavior of both vehicles, in the presence of a lead vehicle and a vehicle in front of it. Therefore, when the vehicle in front of the lead vehicle decelerates, the driver also correspondingly decelerates and enters the railway crossing while keeping larger space headway. However, when the vehicle in front of the lead vehicle does not decelerate, the driver enters the railway crossing while maintaining small space headway. In a road environment where there is an intersection immediately after a railway crossing, a situation in which the vehicle in front of the lead vehicle drives straight ahead but the lead vehicle decelerates and stops for a right or a left turn is likely to occur. In such a road traffic environment, it is highly advisable to alert drivers to slow down and to increase their space headway to a reasonable distance before they enter the railway crossing. This gradual deceleration is expected to increase the space headway to a reasonable distance before the crossing and to thus have less impact on traffic flow than the current obligatory full stop. Analysis of ways to increase space headway and thus maintain smooth traffic flow is an issue for future investigation.

7. Advantage of the New Assessment Method of Driver Behavior

We proposed a new method for predicting the driving behavior in combination with a driving simulator experiment data and a field experiment data. Here, we hypothesized that many data with various conditions could not be collected in the case of the experiments focusing on driver's responses to the incident that happened because the participant became cautious about the succeeding event. In our driving simulator experiments, the drivers participated for three days and they experienced one kind of other vehicles' decelerating condition at total of three times at each speed condition: we collected 90 data under each other vehicles' decelerating condition and at each speed condition. In our collected data, there were no data where the participant did not react to the movement of the leading vehicle (including the release of the accelerator pedal) due to their behavior adaptive to the other vehicle's decelerating behavior before entering the railway crossing.

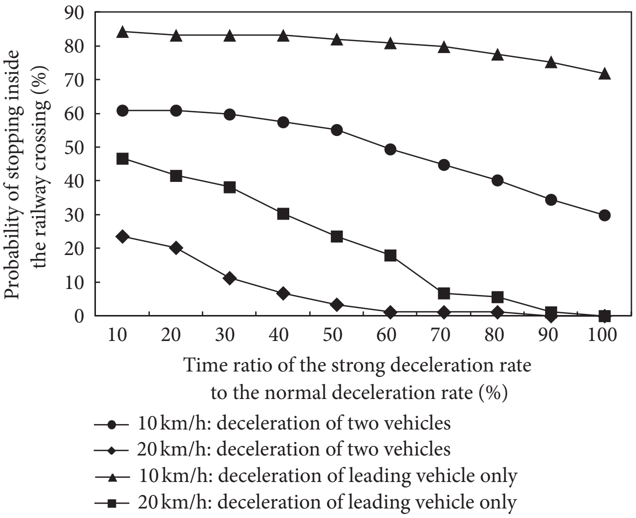


Figure 8: Simulation of the probability of becoming trapped based on time ratio between the driver's strong deceleration rate and the normal deceleration rate.

Another advantage of the new assessment method of driving behavior is that we could simulate the occurrence probability if different behaviors were observed when facing the target situation. For example, we could simulate the influences of the time ratio of driver's deceleration rate on the probability of becoming trapped. The ratio of the time of the strong deceleration rate to the normal deceleration rate was set at 10% during the driver's decelerating behavior in the abovementioned evaluation, which was observed in the field experiment data. Figure 8 presents the changes of the probabilities of entrapment under two kinds of traffic conditions at 10 and 20 km/h according to the changes of the time ratio of the driver's deceleration rate.

The highest probability of becoming trapped was found at the slowest speed even when the driver slowed down with the strong deceleration since the driver first began to brake until they reached a vehicle speed of 0 km/h. This estimation result supports the impact of the small space headway and somewhat longer driver's response time at 10 km/h on the probability of becoming trapped. On the other hand, at 20 km/h when both the lead vehicle and the one in front of it slowed down and when only the lead vehicle slowed down, more than 70% of the time ratio of the strong deceleration while decelerating would lead to the probability of entrapment below 10%, whose result is similar to the probability at 50 km/h. These findings may contribute to developing the presentation timing of the warning to prevent becoming trapped. The time for the driver to recognize that the lead vehicle will stop immediately after driving through the railway crossing can be calculated from the time ratio of the driver's strong deceleration to the normal deceleration. This time can be used to determine the timing for the warning presentation that encourages the driver in the hard deceleration to stop before the railway crossing.

8. Limitations

The deceleration rates for the braking distances in the simulation were calculated based on the deceleration data collected on actual intersections. The deceleration rates while stopping behind a leading vehicle at an intersection were considered to be similar to the deceleration rates while stopping behind a lead vehicle before a railway crossing. However, the estimated braking distances did not include a driver's attitude to the railway crossing, that is, how a driver might assume the railroad track hazard and might react to the presence of the track. Japanese drivers would not enter the crossing unless they have a clear space on the other side under the current traffic law. It is anticipated that the current Japanese drivers' attitude toward entering the track might change as they become familiar with the removal of the obligatory stop, after amending the law to remove this obligation. This driver's reaction to the presence of the railway crossing will be one of the key factors influencing the probability of a driver becoming trapped. This attitude might be relevant to a driver's acceptable range of being involved in the risky condition. Further research will be focused on investigating driver's individual characteristics leading to driver's risky behavior when crossing the track and estimating the impact of these characteristics on the probability of entrapment.

This experiment has examined the impacts of entrance speeds as well as the impacts of the movements of a first and second vehicle in front. In future simulations, we will create a multilane road with a railway crossing and conditions where the lanes next to the driver's lane are flowing smoothly but the lead vehicle in the driver's lane decelerates and stops owing to congestion on the road ahead. We will assess the probability of the driver becoming trapped under these conditions. The railway crossing environment that is currently configured on our driving simulator is based on the premise that a railway crossing should be visually recognizable from far away. In the future, we will focus more on the structure of the road and railway crossing to examine the probability of entrapment at a railway crossing on a visually poorly structured road with, for example, curves and uphill slopes.

9. Conclusions

Our aim was to assess how removing the obligation of drivers to stop before entering railway crossings would affect the probability of their becoming trapped, using our proposed new estimation method of the driving behavior. We examined this by using four vehicle speed conditions (10, 20, 30, and 50 km/h) under the hypothetical condition of a lead vehicle stopping immediately after driving through the railway crossing while following two vehicles.

We obtained the following results.

(i) At vehicle speeds of 10 to 50 km/h, the probability of becoming trapped was the highest at 10 km/h, mainly owing to the small space headway.

(ii) When the vehicle in front of the lead vehicle drove through at a constant speed and the lead vehicle decelerated, there was a greater probability of the test vehicle becoming trapped than when the two vehicles in front both decelerated.

Because the probability of becoming trapped was highest at 10 km/h, the findings suggest that, if the obligation to stop were removed, then the probability of becoming trapped would be lower if drivers were to drive through the crossing while maintaining their approach speed (e.g., driving through at a vehicle speed of 50 km/h if the road speed limit were 50 km/h) rather than decelerating to a very slow speed.

The proposed evaluation method of the driving behavior could be applied to the situations other than entering the railway crossing. For example, driver's emergency avoidance to a sudden appearance of a pedestrian is one of traffic situations potentially leading to the traffic accidents. Once a driver encountered a pedestrian rushing out in front of the driver's vehicle, he/she would become careful of other pedestrians standing along the road. The probability of collisions with a pedestrian suddenly traveling across could be assessed using the simulator data measuring the driver's response to movements of the pedestrian and the field data measuring the driver's deceleration during the emergency avoidance situations on actual roads.

References

[1] A. Kemeny and F. Panerai, "Evaluating perception in driving simulation experiments," *Trends in Cognitive Sciences*, vol. 7, no. 1, pp. 31–37, 2003.

[2] M. Akamatsu, M. Okuwa, and M. Onuki, "Development of hi-fidelity driving simulator for measuring driving behavior," *Journal of Robotics and Mechatronics*, vol. 13, no. 4, pp. 409–418, 2001.

[3] M. Pinto, V. Cavallo, T. Ohlmann, S. Espie, and J. Roge, "The perception of longitudinal accelerations: what factors influence braking manoeuvers in driving simulators?" in *Proceedings of Driving Simulation Conference Europe (DSC '04)*, pp. 139–151, 2004.

[4] E. R. Boer, A. Girshick, T. Yamamura, and N. Kuge, "Experiencing the same road twice: a driver centred comparison between simulation and reality," in *Proceedings of the Driving Simulation Conference (DSC '00)*, pp. 6–8, 2000.

[5] M. Akamatsu and M. Onuki, "Trends in technologies for representing the real world in driving simulator environments," *Review of Automotive Engineering*, vol. 29, no. 4, pp. 611–618, 2008.

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2013, Article ID 924170, 13 pages
http://dx.doi.org/10.1155/2013/924170

Review Article

Development and Evaluation of Automotive Speech Interfaces: Useful Information from the Human Factors and the Related Literature

Victor Ei-Wen Lo[1,2] and Paul A. Green[1]

[1] *Driver Interface Group, University of Michigan Transportation Research Institute, 2901 Baxter Road, Ann Arbor, MI 48109-2150, USA*
[2] *Department of Environmental Health Sciences, School of Public Health, University of Michigan, 1415 Washington Heights, Ann Arbor, MI 48109-2029, USA*

Correspondence should be addressed to Victor Ei-Wen Lo; loe@umich.edu

Received 11 October 2012; Revised 10 January 2013; Accepted 17 January 2013

Academic Editor: Motoyuki Akamatsu

Copyright © 2013 V. E.-W. Lo and P. A. Green. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Drivers often use infotainment systems in motor vehicles, such as systems for navigation, music, and phones. However, operating visual-manual interfaces for these systems can distract drivers. Speech interfaces may be less distracting. To help designing easy-to-use speech interfaces, this paper identifies key speech interfaces (e.g., CHAT, Linguatronic, SYNC, Siri, and Google Voice), their features, and what was learned from evaluating them and other systems. Also included is information on key technical standards (e.g., ISO 9921, ITU P.800) and relevant design guidelines. This paper also describes relevant design and evaluation methods (e.g., Wizard of Oz) and how to make driving studies replicable (e.g., by referencing SAE J2944). Throughout the paper, there is discussion of linguistic terms (e.g., *turn-taking*) and principles (e.g., *Grice's Conversational Maxims*) that provide a basis for describing user-device interactions and errors in evaluations.

1. Introduction

In recent years, automotive and consumer-product manufacturers have incorporated speech interfaces into their products. Published data on the number of vehicles sold with speech interfaces is not readily available, though the numbers appear to be substantial. Speech interfaces are of interest because visual-manual alternatives are distracting, causing drivers to look away from the road, and increasing crash risk. Stutts et al. [1] reported that adjusting and controlling entertainment systems and climate-control systems and using cell phones accounted for 19% of all crashes related to distraction. The fact that the use of entertainment systems is ranked the second among major causes of these crashes arises the argument that speech interfaces should be used for music selection. Tsimhoni et al. [2] reported that 82% less time was needed for drivers to enter an address using a speech interface as opposed to using a keyboard, indicating that a speech interface is preferred for that task. However, using a speech interface still requires cognitive demand, which can interfere with the primary driving task. For example, Lee et al. [3] showed that drivers' reaction time increased by 180 ms when using a complex speech-controlled email system (three levels of menus with four-to-seven options for each menu) in comparison with a simpler alternative (three levels of menus with two options per menu).

Given these advantages, suppliers and automanufacturers have put significant effort into developing speech interfaces for cars. They still have a long way to go. The influential automotive.com website notes the following [13]:

> *In the 2012 . . ., the biggest issue found in today's vehicles are the audio, infotainment, and navigation system's lack of being able to recognize voice commands. This issue was the source of more problems than engine or transmission issues. . . . Over the four years that the survey questions people on voice recognition systems, problems have skyrocketed 137 percent.*

Consumer Reports [14] said the following:

> *I was feeling pretty good when I spotted that little Microsoft badge on the center console. Now I would be able to access all of those cool SYNC features, right? Wrong.*
>
> *When I tried to activate text to speech, I was greeted with a dreadful "Not Supported" display. I racked my brain. Did I do something wrong? After all, my phone was equipped seemingly with every feature known to man. . . . But most importantly, it was powered by Microsoft just like the SYNC system on this 2011 Mustang.*
>
> *Needing guidance, I went to Ford's SYNC website. . . ., I was able to download a 12-page PDF document that listed supported phones. (There is an interactive Sync compatibility guide here, as well.) While I had naively assumed that my high-tech Microsoft phone would work with all the features of "SYNC powered by Microsoft," the document verified that this was not the case. . . . Text to speech would only work with a small handful of "dumbphones" that aren't very popular anymore. Anyone remember the Motorola Razr? That phone was pretty cool a couple of years ago.*

One consumer, in commenting about the Chrysler UConnect system said the following [15]:

> *I have a problem with Uconnect telephone. I input my voice tags but when I then say "Call Mary" the system either defaults to my 'phone book folder or I get 4–6 names on the screen and am asked to "select a line". I should just say "call Mary home" then I should here my voice with calling "Mary home is that correct". Can you assist?*

Thus, it is appropriate to ask what is known now about the design and evaluation of speech interfaces for cars and how they can be improved. Most engineered systems rely on models, equations, and data to predict system performance and evaluate system alternatives early in development. They do not exist for speech interfaces. Thus, for speech interfaces, the emphasis has been on usability testing, often conducted when development is nearly complete and changes are costly to make.

To be more specific, this paper summarizes the state of the art relating to speech interface design in general, as well as a particular simulation model, namely, one to predict user performance when interacting with a speech interface for destination entry and music selection. The model to be developed will allow for exploration of multiple alternative architectures, recognition rates, and command sets, matters that are very expensive to explore experimentally.

The questions addressed in this paper are as follows.

(1) What are some examples of automotive speech interfaces?

(2) Who uses speech interfaces, for what, and how often?

(3) What are the key research results of the user performance using speech interfaces compared with the user performance using visual-manual interfaces?

(4) How should speech interfaces be designed? What are the key design standards and references, design principles, and results from research?

(5) How should speech interfaces be assessed and what should be measured?

To answer these questions, both the human factors literature and the linguistics terms and principles are summarized.

2. What Are Some Examples of Automotive Speech Interfaces?

In the USA, current speech interfaces include Ford SYNC, Chrysler UConnect, GM MyLink, Hyundai Genesis, and Toyota navigation with Entune. The commonly supported applications are navigation (e.g., destination entry, route guidance, and traffic information) and music selection (selecting, playing, and pausing songs on MP3 players, AM/FM/XM radios), as well as those related to cellular phones (answering and placing calls, searching contact lists, and various tasks associated with text messages).

To help the reader understand these applications, some examples follow. The extent to which the user and machine are verbose varies between examples.

⟨Navigation⟩

M: *Command please.*
U: *Find address.*
M: *Find address. The city please.* ⟨Default State is Michigan⟩
U: *Michigan City.*
M: *Please select the respective line or start spelling.*
U: *M I C H I G A N.*
M: *Spell the name in blocks of letters or enter your destination again with change state or enter city.*
U: *Change state.*
M: *Please enter the state name.*
U: *Indiana.*
M: *Please select the respective line or start spelling.*
U: *Line one.*
M: *Line one. Indiana. The city please.*
U: *Michigan City.*
M :*Michigan City. What is the street?*
U: *Wabash Street.*
M: *Wabash Street. And house number or if you do not know that please say show map or start guidance.*
U: *Six zero one.*
M: *Six zero one. Show map or start guidance.*
U: *Start guidance*
M: *Start guidance.*

Table 1: Examples of well-known automotive speech interfaces and the applications they support.

System	Research/product	Navigation	Restaurant finder	Music selection	Audio + CD	Car information	Traffic information	Cellular phone	Text message
CHAT [4, 5]	Research	X	X	X					
CU Move [6]	Research	X					X		
Ford Model U [7]	Research	X		X		X		X	
Linguatronic [8]	Product	X		X		X		X	
SENECA [9]	Research	X			X			X	
SYNC [10]	Product	X		X		X		X	X
VOIC [11]	Research	X		X		X	X		
Volkswagen [12]	Product	X						X	

⟨Restaurant Finder⟩ [4]

U: *Can you find me a Chinese restaurant in Palo Alto with good service?*

M: *There are 11 Chinese restaurants in Palo Alto that have good service. Here are the first few: Little Garden Restaurant China Delight Mandarin Gourmet Restaurant. Should I continue?*

U: *Hmm…Tell me more about, Garden Restaurant?*

M: *Little Garden Restaurant is an expensive Chinese Restaurant. Dress Code is casual, service is good and rating is three stars. It is on El Camino Real in Palo Alto.*

U: *Switch to navigation.*

M: *Take me to little Garden Restaurant, but stop by Bosch Page Mill Office.*

…

⟨Music Selection⟩

M: *Please say a command.*
U: *USB.*
M: *USB. Is it correct?*
U: *Yes.*
M: *USB. Please say a command.*
U: *PLAY TRACK "Like a Prayer."*
M: *Play track "Like a prayer."*

⟨Phone call⟩

M: *Main menu.*
U: *Phone.*
M: *Phone. Please say a command.*
U: *Call XXX XXX XXXX.*
M: *Calling XXX XXX XXXX.*

⟨Radio⟩

U: *AM one thousand six hundred.*

[It did not work if the user said "AM Sixteen Hundred" for this interface.]

M: *AM one thousand six hundred.*

These systems were developed based on ideas from a number of predecessor systems (Tables 1 and 2). Notice that the core functions were navigation, music selection, and cellular phone support, and that many of them started out as either university or collaborative research projects involving several partners. In several cases, the result was either a product or ideas that later led to products. Of them, probably SYNC has received most of the attention.

The CHAT system uses an event-based, message-oriented system for the architecture with core modules of Natural Language Understanding (NLU), Dialogue Manager (DM), Content Optimization (CO), Knowledge Management (KM), and Natural Language Generation (NLG). CHAT uses the Nuance 8.5 speech recognition engine with class-based n-grams and dynamic grammars, and Nuance Vocalizer as the Text-to-Speech engine. There are three main applications—navigation, MP3 music player, and restaurant finder—to represent important applications in a vehicle [4, 5]. The example for restaurant finder shown earlier is a CHAT dialog.

CU-Move system is an in-vehicle, naturally spoken dialogue system, which can get real-time navigation and route-planning information [6]. The dialogue system is based on the MIT Galaxy-II Hub architecture with base components from CU-Communication system, which is mixed initiative and event driven. This system automatically retrieves the driving direction through Internet with route provider. The dialogue system uses the CMU Sphinx-II speech recognizer for speech recognition and Phoenix Parser for semantic parsing.

A prototype of a conversation system was implemented on the Ford Model U Concept Vehicle and was first shown in 2003 [7]. This system is used for controlling several noncritical automobile operations using speech recognition and a touch screen. The speech recognizer used in this system was speech2Go with adapted acoustic model and other enhancements to improve the speech accuracy. The dialogue manager was a multimodal version of the ETUDE, described by a recursive transition network. Supported applications were climate control, telephone, navigation, entertainment, and system preferences.

Linguatronic is a speech-based command and control system for telephone, navigation, radio, tape, CD, and other applications. The recognizer used in this device was speaker-independent [8].

Table 2: Origins of some well-known automotive speech applications.

System	Full name	Developed by
CHAT	Conversational Helper for Automotive Tasks	Center for the Study of Language and Information at Stanford University, Research and Technology Center at Bosch, Electronics Research Lab at Volkswagen of America, and Speech Technology and Research Lab at SRI International [4, 5].
CU Move	Colorado University Move	University of Colorado speech group in 1999 [6].
Ford Model U		Ford [7].
Linguatronic		DaimlerChrysler Research and Technology in Ulm, Germany and TEMIC in 1996 [8].
SENECA	Speech control modules for Entertainment, Navigation, and communication Equipment in CArs	EU-project involving DaimlerChrysler, TEMIC Research and Department of Information Technology, University of Ulm [9].
SYNC		Ford in collaboration with Microsoft and Nuance [10].
VOIC	Virtual Intelligent Co-Driver	European project funded by five different partners: Robert Bosch GmbH, DaimlerChrysler AG, ITCirst, the University of Southern Denmark, and Phonetic Topographics N.V [11].
Volkswagen		Volkswagen [12].

SENECA SLDS consists of five units: COMMAND head unit connected via an optical Domestic Digital Bus to the Global System for Mobile Communication module, the CD Changer, and Digital Signal Processing module [9]. The system is a command-based speech control of entertainment (radio and CD), navigation, and cellular phones. The speech recognition technology of SENECA SLDS is based on the standard Linguatronic system using the following methods to match the user speech: spell matcher, Java Speech Grammar Format, voice enrollments (user-trained words), and text enrolments. For the dialogue processing, the ENECA SLDS uses a menu-based Command & Control dialogue strategy, including top-down access for main function and side access for subfunction.

SYNC is a fully integrated, voice-activated in-vehicle communication and entertainment system [10] for Ford, Lincoln, and Mercury vehicles in North America. Using commands in multiple languages, such as English, French or Spanish, drivers can operate navigation, portable digital music players, and Bluetooth-enabled mobile phones. The example for music selection shown earlier is a SYNC dialog.

VICO was a research project that concerned a natural-language dialogue prototype [11]. As the interface did not exist, researchers used the Wizard of Oz method to collect the human-computer interaction data. Here, a human operator, the wizard, was simulated system components—speech recognition, natural language understanding, dialogue modeling, and response generation. The goal of this project was to develop a natural language interface allowing drivers to get time, travel (navigation, tourist attraction, and hotel reservation), car, and traffic information safely while driving.

Volkswagen also developed its own in-vehicle speech system [12]. Detailed information about the architecture and methods used to design the system are not available. Supported applications include navigation and cellular phones.

The best-known nonautomotive natural speech interface is Siri, released by Apple in October 2011. Siri can help users make a phone call, find a business and get directions, schedule reminders and meetings, search the web, and perform other tasks supported by built-in apps on the Apple iPhone 4S and iPhone 5.

Similarly, Google's Voice Actions supports voice search on Android phones (http://www.google.com/mobile/voice-actions/, retrieved May 14, 2012). This application supports sending text messages and email, writing notes, calling businesses and contacts, listening to music, getting directions, viewing a map, viewing websites, and searching webpages. Both Siri and Voice Actions require off-board processing, which is not the case for most in-vehicle speech interfaces.

3. Who Uses Speech Interfaces, for What, and How Often?

Real-world data on the use of speech applications in motor vehicles is extremely limited. One could assume that anyone who drives is a candidate user, but one might speculate that the most technically savvy are the most likely users.

How often these interfaces are used for various tasks is largely unknown. The authors do not know of any published studies on the frequency of use of automotive speech interfaces by average drivers, though they probably exist.

The most relevant information available is a study by Lo et al. [28] concerning navigation-system use, which primarily concerned visual-manual interfaces. In this study, 30 ordinary drivers and 11 auto experts (mostly engineers employed by Nissan) completed a survey and allowed the authors to download data from their personal navigation systems. Data was collected regarding the purpose of trips (business was most common) and the driver's familiarity with the destination. Interestingly, navigation systems were used to drive to familiar destinations. Within these two groups, use of speech interfaces was quite limited, with only two of the ordinary drivers and two of the auto experts using speech interfaces.

Table 3: Speech interface performance statistics from selected bench-top studies.

System	CHAT [4]	CHAT [5]	CU Communicator [16]	CU Move [6]	SENECA [9]	VOIC [11]	Volkswagen [12]
Tasks	(1) NAV (2) Restaurant finder (RF)	(1) MP3 (2) Restaurant finder (RF)	Phone for travel plan	NAV	(1) NAV (2) Phone dialing (3) Address book	(1) NAV (2) Current time (3) Tourism (4) Fuel (5) Car manual (6) Hotel reservation (7) Traffic information	(1) NAV (2) Map control (3) Phone
Completion time (s)			260.3		Average: 63	(1) 73 (2) 15 (3) 146 (4) 78 (5) 57 (6) 180 (7) 53	
Completion rate	98%	MP3: 98% RF: 94%	73.6%		Average: 79%		
Turns[1]	2.3[1]	RF: 4.1[1]	User: 19 Machine: 19				
Word recognition Accuracy[2] (%)	NAV: 85.5% RF: 85%	MP3: 90% RF: 85%					
Word error rate (%)			26%	30–65%			
User satisfaction rating[3]	1.98	MP3: 2.24 RF: 2.04					

[1] *Turn* is defined as one user utterance to the system during a dialog exchange between the user and the system while attempting to perform the task.

[2] Word recognition accuracy (WA).

WA = 100(1 − (Ws + Wi + Wd)/W)%.

W: total number of words in reference.

Ws: number of reference words which were substituted in output.

Wi: number of reference words which were inserted in output.

Wd: number of reference words which were deleted in output.

[3] User satisfaction rating: 1 = strong agreement; 5 = strong disagreement.

Table 4: Driving performance statistics from selected studies (S: speech; M: manual, K: keyboard).

Study	Method	Lane keeping	Brake reaction time	Peripheral detection time	Following distance
Carter and Graham [17]	Simulator	S < M	S < M		
Forlines et al. [18]	Simulator	S < M	No diff.		
Garay-Vega et al. [19]	Simulator	No diff.			
Gärtner et al. [20]	On road	S < M			
Itoh et al. [21]	Simulator	S < M	No diff.		
Maciej and Vollrath [22]	Simulator	S < M			
McCallum et al. [23]	Simulator	No diff.	No diff.		
Minker et al. [9]	On road	S < M			
Ranney et al. [24]	On road			S < M (0.8 versus 0.87 s)	
Shutko et al. [25]	Simulator	S < M		S < M (Except incoming call)	
Tsimhoni et al. [26]	Simulator	S < K			S < K (88 versus 167 m)
Villing et al. [27]	On road				

Table 5: Task performance statistics from selected studies (S: speech; M: manual).

Study	Task completion time	Speech recognizer rate	Task completion rate
Carter and Graham [17]	S > M	92.7%	
Forlines et al. [18]	S < M (18.0 versus 25.2 s)		
Garay-Vega et al. [19]	S (dialog-based) > M S (query-based) < M		
Gärtner et al. [20]	S > M Simple: 24.6 versus 12.8 s Complex: 74.4 versus 58.7 s	79.4% (recognition error rate: 20.6%)	
Minker et al. [9]	S < M (63 versus 84 sec)		S < M (79 versus 90%)
Ranney et al. [24]	No difference		
Shutko et al. [25]	S < M (except dialing phone)		
Villing et al. [27]	S > M		

The paper also contains considerable details on the method of address entry (street address being used about half of the time followed by point of interest POI) and other information useful in developing evaluations of navigation systems.

Also relevant is the Winter et al. [29] data on typical utterance patterns for speech interfaces, what drivers would naturally say if unconstrained. Included in that paper is information on the number and types of words in utterances, the frequency of specific words, and other information needed to recognize driver utterances for radio tuning, music selection, phone dialing, and POI and street-address entry. Takeda et al. [30] present related research on in-vehicle corpora, which may be a useful resource to address on who, when, and how often the driver used the speech interfaces.

4. What Are the Key Research Results of the User Performance Using Speech Interfaces Compared with the User Performance Using Visual-Manual Interfaces?

There have been a number of studies on this topic. Readers interested in the research should read Barón and Green [31] and then read more recent studies.

Using Barón and Green [31] as a starting point, studies of the effects of speech interfaces on driving are summarized in four tables. Table 3 summarizes bench-top studies of various in-vehicle speech interfaces. Notice that the value of the statistics varied quite widely between speech interfaces, mainly because the tasks examined were quite different. As an example for CU-Communicator [16], the task required the subject to reserve a one-way or round-trip flight within or outside the United States with a phone call. Performing this task involved many turns between users and machines (total 38 turns) and the task took almost 4.5 minutes to complete. Within speech interfaces, task-completion time varied from task to task depending on the task complexity [11, 12].

Table 4, which concerns driving performance, shows that the use of speech interfaces as opposed to visual-manual interfaces led to better lane keeping (e.g., lower standard deviation of lane position).

Table 5 shows that task completion times for speech interfaces were sometimes shorter than that for visual-manual interfaces and sometimes longer, even though people speak faster than they can key in responses. This difference is due to the inability of the speech interface to correctly recognize what the driver says, requiring utterances to be repeated. Speech recognition accuracy was an important factor that

Table 6: Subjective rating and driving behavior statistics from selected studies (S: speech; M: manual).

Study	Subjective rating		Driving behavior—glances	
	Workload	Preference	Glance duration	Number of glances
Carter and Graham [17]	S < M			
Faerber and Meier-Arendt [32]		S is preferred	Radio or CD control: M: 1 s; S: 1/3 s Using phone: M: 1 s	Radio or CD control: M: 3; S: 1.1 Using phone: M: 12; S: 0
Garay-Vega et al. [19]	S (query-based) < M		S < M	
Gärtner et al. [20]	No difference			Simple task: no difference Complex tasks: S < M
Itoh et al. [21]	S < M		S < M	
Maciej and Vollrath [22]			S < M	
McCallum et al. [23]	S < M			
Minker et al. [9]		S is preferred		
Shutko et al. [25]			S < M (Except receiving an incoming call)	

affected the task performance. Kun et al. [33] reported that low recognition accuracy (44%) can lead to greater steering angle variance. Gellatly and Dingus [34] reported that driving performance (peak lateral acceleration and peak longitudinal acceleration) was not statistically affected until the 60% recognition accuracy level was reached. Gellatly and Dingus [34] also showed that the task completion time was also affected when the speech recognition accuracy was lower than 90%. Although speech recognition accuracy was found to affect driving and task performance, no research has been reported on drivers' responses to errors, how long drivers need to take to correct errors, or what strategies drivers use to correct errors. Understanding how users interact with the spoken dialogue systems can help designers improve system performance and make drivers feel more comfortable using speech interfaces.

Table 6 shows that when using speech interfaces while driving, as opposed to visual-manual interfaces, subjective workload was less, fewer glances were required, and glance durations were shorter.

In general, driving performance while using speech interfaces is generally better than when using visual-manual interfaces. That is, speech interfaces are less distracting.

5. How Should Speech Interfaces Be Designed? What Are the Key Design Standards and References, Design Principles, and Results from Research?

5.1. Relevant Design and Evaluation Standards. For speech interfaces, the classic design guidelines are that of Schumacher et al. [35], and the one set that is not very well known, but extremely useful, is the Intuity guidelines [36]. Najjar et al. [37] described user-interface design guidelines for speech recognition applications. Hua and Ng [38] also proposed guidelines on in-vehicle speech interfaces based on a case study.

Several technical standards address the topic of the evaluation of speech system performance. These standards, such as ISO 9921: 2003 (Ergonomics—Assessment of speech communication), ISO 19358: 2002 (Ergonomics—Construction and application of tests for speech technology), ISO/IEC 2382-29: 1999 (Artificial intelligence—Speech recognition and synthesis), and ISO 8253-3: 2012 (Acoustics—Audiometric tests methods—Part 3: Speech Audiometry), focus on the evaluation of the whole system and its components [39–42]. However, no usability standards related to speech interfaces have emerged other than ISO/TR 16982: 2002 (Ergonomics of human-system interaction—Usability methods supporting human-centered design) [43].

From its title (Road vehicles—Ergonomic aspects of transport information and control systems—Specifications for in-vehicle auditory presentation), one would think that ISO 15006: 2011 [44] is relevant. In fact, ISO 15006 concerns nonspoken warnings.

There are standards in development. SAE J2988, Voice User Interface Principles and Guidelines [45], contains 19 high-level principles (e.g., principle 17: "Audible lists should be limited in length and content so as not to overwhelm the user's short-term memory."). Unfortunately, no quantitative specifications are provided. The draft mixes definitions and guidance in multiple sections making the document difficult to use, does not support guidance with references, and, in fact, has no references.

The National Highway Traffic Safety Administration (NHTSA) of the US Department of Transportation posted proposed visual-manual driver-distraction guidelines for in-vehicle electronic devices for public comment on February 15, 2012 (http://www.nhtsa.gov/About+NHTSA/Press+Releases/2012/U.S.+Department+of+Transportation+Proposes+'Distraction'+Guidelines+for+Automakers, retrieved May 15, 2012). NHTSA has plans for guidelines for speech interfaces.

The distraction focus group of the International Telecommunication Union (FG-Distraction-ITU) is interested in speech interfaces and may eventually issue documents on this topic, but what and when are unknown. In addition, various ITU documents that concern speech-quality assessment may be relevant, though they were intended for telephone applications. ITU-P.800 (methods for subjective determination of transmission quality) and related documents are of particular interest. See http://www.itu.int/rec/T-REC-P/e/.

5.2. Key Books. There are a number of books on speech interface design, with the primary references being Hopper's classic [46], Balentine and Morgan [47], Cohen et al. [48], and Harris [49]. A more recent reference is Lewis [50].

5.3. Key Linguistic Principles. The linguistic literature provides a framework for describing the interaction, the kinds of errors that occur, and how they could be corrected. Four topics are touched upon here.

5.3.1. Turn and Turn-Taking. When can the user speak? When does the user expect the system to speak? Taking a turn refers to an uninterrupted speech sequence. Thus, the back-and-forth dialog between a person and a device is turn-taking, and the number of turns is a key measure of an interface's usability, with fewer turns indicating a better interface. In general, overlapping turns, where both parties speak at the same time, account for less than 5% of the turns that occur while talking [51]. The amount of time between turns is quite small, generally less than a few hundred milliseconds. Given the time required to plan an utterance, planning starts before the previous speaker finishes the utterance.

One of the important differences between human-human and human-machine interactions is that humans often provide nonverbal feedback that indicates whether they understand what is said (e.g., head nodding), which facilitates interaction and control of turn-taking. Most speech interfaces do not have the ability to process or provide this type of feedback.

A related point is that most human-human interactions accept interruptions (also known as barge-in), which makes interactions more efficient and alters turn taking. Many speech interfaces do support barge-in, which requires the users to press the voice-activation button. However, less than 10% of subjects (unpublished data from the authors) knew and used this function.

5.3.2. Utterance Types (Speech Acts). Speech acts refer to the kinds of utterances made and their effect [53]. According to Akmajian et al. [54], there are four categories of speech acts.

(i) Utterance acts include uttering sounds, syllables, words, phrases, and sentences from a language including filler words ("umm").

(ii) Illocutionary acts include asking, promising, answering, and reporting. Most of what is said in a typical conversation is this type of act.

(iii) Perlocutionary acts are utterances that produce an effect on the listener, such as inspiration and persuasion.

(iv) Propositional acts are acts in which the speaker refers to or predicts something.

Searle [55] classifies speech acts into five categories.

(i) Assertives commit the speaker to address something (suggesting, swearing, and concluding).

(ii) Directives get the listener to do something (asking, ordering, inviting).

(iii) Commissives commit the speaker to some future course of action (promising, planning).

(iv) Expressives express the psychological state of the speaker (thanking, apologizing, welcoming).

(v) Declarations bring a different state to either speaker or listener (such as "You are fired").

5.3.3. Intent and Common Understanding (Conversational Implicatures and Grounding). Sometimes speakers can communicate more than what is uttered. Grice [56] proposed that conversations are governed by the cooperative principle, which means that speakers make conversational contributions at each turn to achieve the purpose or direction of a conversation. He proposed four high levels conversational maxims that may be thought of as usability principles (Table 7).

5.3.4. Which Kinds of Errors Can Occur? Skanztze [52] provides one of the best-known schemes for classifying errors (Table 8). Notice that Skanztze does so from the perspective of a device presenting an utterance and then processing a response from a user.

Véronis [57] presents a more detailed error-classification scheme that considers device and user errors, as well as the linguistic level (lexical, syntactic, semantic). Table 9 is an enhanced version of that scheme. Competence, one of the characteristics in his scheme, is the knowledge the user has of his or her language, whereas performance is the actual use of the language in real-life situations [58]. Competence errors result from the failure to abide by linguistic rules or from a lack of knowledge of those rules ("the information from users is not in the database"), whereas performance errors are made despite the knowledge of rules ("the interface does not hear users' input correctly").

As an example, a POI category requested by the user that was not in the database would be a semantic competence error. Problems in spelling a word would be a lexical performance error. Inserting an extra word in a sequence ("iPod iPod play . . .") would be a lexical performance error.

A well-designed speech interface should help avoiding errors, and, when they occur, facilitating correction. Strategies to correct errors include repeating and rephrasing the utterances, spelling out words, contradicting a system response, correcting using a different modality (e.g., manual entry instead of speech), and restarting, among others [59–62].

Knowing how often these strategies occur suggests what needs to be supported by the interface. The SENECA project [9, 20] revealed that the most frequent errors for navigation tasks were spelling problems of various types, entering or

Table 7: Grice's conversational maxims (with examples added by the authors).

Maxim	Example	Guidance
Maxim of Quantity: be informative.	Machine: Please say the street name. User: 2901 Baxter Road (2901 is the house number)	(i) Make your contribution of information as is required, that is, for the current purpose of the conversation. (ii) Do not make your contribution more informative than is required.
Maxim of Quality: make your contribution one that is true:	U: "Toledo Zoo, Michigan" (but Toledo is in Ohio)	(i) Do not say what you believe to be false. (ii) Do not say that for which you lack evidence.
Maxim of Relevance: be relevant.	U: "I want to go to "Best Buy"" and the system responds with all Best Buy stores, including those hundreds of miles away, not just the local ones.	
Maxim of Manner: be perspicuous	(i) M: Please say set as destination, dial or back. U: Dial. O no Don't dial, Back (User want to say "back") (ii) M: Please say the POI category. U: Let's see. Recreation	(i) Avoid obscurity of expression (ii) Avoid unnecessary ambiguity (iii) Be brief (avoid unnecessary prolixity) (iv) Be orderly.

Table 8: Examples of errors in different modules of speech-controlled interfaces (adapted from Skanztze [52]).

Modules	Possible sources of errors
Speech detection	Truncated utterances, artifacts such as noise and side talk; barge-in problems
Speech recognition	Insertions, deletions, substitutions
Language processing/parsing	Concept failure, speech act tagging
Dialogue manager	Error in reference resolution, error in plan recognition
Response generation	Ambiguous references, too much information presented at once, TTS quality, audio quality

Table 9: Enhanced version of Véronis [57] error-classification scheme.

	System		User	
	Performance	Competence	Performance	Competence
Lexical level (word)	Letter substitution Letter insertion Letter deletion	Word missing in dictionary Missing inflection rule	Letter substitution Letter insertion Letter deletion Letter transposition Syllabic error Slips of tongue	Nonword or completely garbled word
Syntactic level (sentence structure)		Missing rule	Word substitution Word insertion Word deletion Word transposition	Construction error
Semantic level (meaning)		Incomplete or contradictory knowledge representation Unexpected situation		*Conceptual error* includes: incomplete or contradictory knowledge representation *Pragmatic error* includes: dialogue law violation

choosing the wrong street, and using wrong commands. For phone dialing tasks, the most frequent errors were stops within digit sequences. In general, most of the user errors were vocabulary errors (partly spelling errors), dialogue flow errors, and PTA (push to active) errors, that is, missing or inappropriate PTA activation.

Lo et al. [63] reported that construction and relationship errors were 16% and 37%, respectively. Construction errors occur when subjects repeat words, forget to say command words (a violation of grounding), or forget to say any other words that were given. Relationship errors occur when subjects make incorrect matches between the given words

Table 10: Variables used for evaluating entire systems or system modules.

Module	Variables
Whole system	Task completion time, task completion rate, transaction success, number of interaction problems, query density, concept efficiency
Speech recognition	Word and sentence error rate, vocabulary coverage, perplexity
Speech synthesizer	User perception, speech intelligibility, pleasantness, naturalness
Language understanding	Lexical coverage, grammar coverage, real-time performance, concept accuracy, concept error rate

and song title, album name, and/or artist name. Relationship errors were common because subjects were not familiar with the given songs/albums/artists.

6. How Should Speech Interfaces Be Assessed and What Should Be Measured?

6.1. What Methods Should Be Used? Given the lack of models to predict user performance with speech interfaces, the evaluation of the safety and usability (usability testing) of those interfaces has become even more important. Evaluations may either be performed only with the system itself (on a bench top) or with the system integrated into a motor vehicle (or a simulator cab) while driving.

The most commonly used method to evaluate in-vehicle speech interfaces is the Wizard of Oz method [4, 5, 11, 16, 64–66], sometimes implemented using Suede [67]. In a Wizard of Oz experiment, subjects believe that they are interacting with a computer system, not a person simulating one. The "wizard" (experimenter), who is remote from the subject, observes the subject's actions and simulates the system's responses in real-time. To simulate a speech-recognition application, the wizard would type what users say, or in a text-to-speech system, they read the text output, often in a machine-like voice. Usually, it is much easier to tell a person how to emulate a machine than to write the software to tell a computer to do it. The Wizard of Oz method allows for the rapid simulation of speech interfaces and the collection of data from users interacting with a speech interface, allowing for multiple iterations of the interface to be tested and redesigned.

6.2. What Should Be Measured? Dybkjaer has written several papers on speech interface evaluation, the most thorough of which is Dybkjær et al. [68]. That paper identified a number of variables that could be measured (Table 10), in part because there are many attributes to consider.

Walker et al. [69] proposed a framework of usability evaluation of spoken dialogue systems, known as PARADISE (PARAdigm for DIalogue System Evaluation). (See [70] for criticisms.) Equations were developed to predict dialog efficiency (which depends on mean elapsed time and the mean number of user moves) and dialog quality costs (which depends on the number of missing responses, the number of errors, and many other factors, and task success, measured by the Kappa coefficient and defined below):

$$\kappa = \frac{(P(A) - P(E))}{1 - P(E)}, \qquad (1)$$

where $P(A)$ = proportion of times that the actual set of dialogues agree with scenario keys; $P(E)$ = proportion of times that the dialogues and the keys are expected to agree by chance.

In terms of performance while driving, there is no standard or common method for evaluating speech interfaces, with evidence from bench-top, simulator, and on-road experiments being used. There are two important points to keep in mind when conducting such evaluations. First, in simulator and on-road experiments, the performance on the secondary speech interface task depends on the demand of the primary driving task. However, the demand or workload of that task is rarely quantified [71, 72]. Second, there is great inconsistency in how secondary-task performance measures are defined, if they are defined at all, making the comparison of evaluations quite difficult [73]. (See [74] for more information.) Using the definitions in SAE Recommended Practice J2944 [75] is recommended.

7. Summary

The issues discussed in this paper are probably just a few of those which should be considered in a systematic approach to the design and development of speech interfaces.

7.1. What Are Some Examples of Automotive Speech Interfaces? Common automotive examples include CHAT, CU Move, Ford Model U, Linguatronic, SENECA, SYNC, VOIC, and Volkswagen. Many of these examples began as collaborative projects that eventually became products. SYNC is the best known.

Also important are nonautomotive-specific interfaces that will see in-vehicle use, in particular, Apple Siri for the iPhone and Google Voice Actions for Android phones.

7.2. Who Uses Speech Interfaces, for What, and How Often? Unfortunately, published data on who uses speech interfaces and how real drivers in real vehicles use them is almost zero. There are several studies that examine how these systems are used in driving simulators, but those data do not address this question.

7.3. What Are the Key Research Results of the User Performance Using Speech Interfaces Compared with the User Performance Using Visual Manual Interfaces? To understand the underlying research, Barón and Green's study [31] is a recommended summary. Due to the difference of task complexity while testing, comparing alternative speech systems is not so easy.

However, when compared with visual-manual interfaces, speech interfaces led to consistently better lane keeping, shorter peripheral detection time, lower workload ratings, and shorter glance durations away from the road. Task completion time was sometimes greater and sometimes less, depending upon the study.

7.4. How Should Speech Interfaces Be Designed? What Are the Key Design Standards and References, Design Principles, and Results from Research? There are a large number of relevant technical standards to help guide speech interfaces. In terms of standards, various ISO standards (e.g., ISO 9921, ISO 19358, ISO 8253) focus on the assessment of the speech interaction, not on design. Speech-quality assessment is considered by ITU-P.800. For design, key guidelines include [35–38]. A number of books also provide useful design guidance including [46–50].

Finally, the authors would recommend that any individual seriously engaged in speech-interface design should understand the linguistic terms and principles (*turns*, *speech acts*, *grounding*, etc.) as the literature provides several useful frameworks for classifying errors and information that provides clues as to how to reduce errors associated with using a speech interface.

7.5. How Should Speech Interfaces Be Assessed and What Should Be Measured? The Wizard of Oz method is commonly used in the early stages of interface development. In that method, an unseen experimenter behind the scenes simulates the behavior of a speech interface by recognizing what the user says or is speaking in response to what the user says, or both. Wizard of Oz simulations take much less time to implement than other methods.

As automotive speech interfaces move close to production, the safety and usability of those interfaces are usually assessed in a driving simulator, and sometimes on the road. The linguistics literature provides a long list of potential measures of the speech interface that could be used, with task time being the most important. Driving-performance measures, such as standard deviation of lane position and gap variability, are measured as eyes-off-the-road time. These studies often have two key weaknesses: (1) the demand/workload of the primary task is not quantified, yet performance on the secondary speech task can depend on its demand and (2) measures and statistics describing primary task performance are not defined. A solution to the first problem is to use equations being developed by the second author to quantify primary task workload. The solution to the second problem is to use the measures and statistics in SAE Recommended Practice J2944 [75] and refer to it.

Driver distraction is and continues to be a major concern. Some view speech interfaces as a distraction-reducing alternative to visual-manual interfaces. Unfortunately, at this point, actual use by drivers and data on that use is almost zero. There is some information on how to test speech interfaces, but technical standards cover only a limited number of aspects.

There is very little to support design other than guidelines. For most engineered systems, developers use equations and models to predict system and user performance, with testing serving as verification of the design. For speech interfaces, those models do not exist. This paper provides some of the background information needed to create those models.

References

[1] J. C. Stutts, D. W. Reinfurt, and L. Staplin, "The role of driverdistraction in traffic crashes," AAA Foundation for Traffic Safety, Washington, DC, USA, 2001, https://www.aaafoundation.org/sites/default/files/distraction%20%281%29.pdf.

[2] O. Tsimhoni, D. Smith, and P. Green, "Address entry while driving: speech recognition versus a touch-screen keyboard," *Human Factors*, vol. 46, no. 4, pp. 600–610, 2004.

[3] J. D. Lee, B. Caven, S. Haake, and T. L. Brown, "Speech-based interaction with in-vehicle computers: the effect of speech-based e-mail on drivers' attention to the roadway," *Human Factors*, vol. 43, no. 4, pp. 631–640, 2001.

[4] F. Weng, B. Yan, Z. Feng et al., "CHAT to your destination," in *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*, pp. 79–86, Antwerp, Belgium, 2007.

[5] F. Weng, S. Varges, B. Raghunathan et al., "CHAT: a conversational helper for automotive tasks," in *Proceedings of the 9th International Conference on Spoken Language Processing (InterSpeech/ICSLP '06)*, pp. 1061–1064, Pittsburgh, Pa, USA, September 2006.

[6] J. H. L. Hansen, J. Plucienkowski, S. Gallant, B. Pellom, and W. Ward, "'CU-Move': robust speech processing for in-vehicle speech system," in *Proceedings of the International Conference on Spoken Language Processing (ICSLP '00)*, vol. 1, pp. 524–527, Beijing, China, 2000.

[7] R. Pieraccini, K. Dayanidhi, J. Bloom et al., "Multimodal conversational systems for automobiles," *Communications of the ACM*, vol. 47, no. 1, pp. 47–49, 2004.

[8] P. Heisterkamp, "Linguatronic product-level speech system for Mercedes-Benz cars," in *Proceedings of the 1st International Conference on Human Language Technology Research*, pp. 1–2, Association for Computational Linguistics, San Diego, Calif, USA, 2001.

[9] W. Minker, U. Haiber, P. Heisterkamp, and S. Scheible, "The SENECA spoken language dialogue system," *Speech Communication*, vol. 43, no. 1-2, pp. 89–102, 2004.

[10] Sync, http://www.ford.com/syncmyride/#/home/.

[11] P. Geutner, F. Steffens, and D. Manstetten, "Design of the VICO spoken dialogue system: evaluation of user expectations by wizard-of-oz experiments," in *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC '02)*, Las Palmas, Spain, 2002.

[12] J. C. Chang, A. Lien, B. Lathrop, and H. Hees, "Usability evaluation of a Volkswagen Group in-vehicle speech system," in *Proceedings of the 1st International Conference on Automotive User Interfaces and Interactive Vehicular Applications (ACM '09)*, pp. 137–144, Essen, Germany, September 2009.

[13] T. Dorchies, "Come again? Vehicle voice recognition biggest problem in J.D. Power and Associates study," http://news.consumerreports.org/cars/2010/08/the-ford-sync-system-reads-text-messages-in-theory-mustang.html.

[14] "Consumer Reports, The Ford SYNC system read test messages…in theory," http://news.consumerreports.org/cars/2010/08/the-ford-sync-system-reads-text-messages-in-theory-mustang.html.

[15] JustAnswer, "Chrysler 300c I have a problems with Uconnect telephone," http://www.justanswer.com/chrysler/51ssy-chrysler-300c-problem-uconnect-telephone.html.

[16] B. Pellom, W. Ward, J. Hansen et al., "University of Colorado dialog systems for travel and navigation," in *Proceedings of the 1st International Conference on Human language Technology Research*, Association for Computational Linguistics, San Diego, Calif, USA, 2001.

[17] C. Carter and R. Graham, "Experimental comparison of manual and voice controls for the operation of in-vehicle systems," in *Proceedings of the 14th Triennial Congress of the International Ergonomics Association and 44th Annual Meeting of the Human Factors and Ergonomics Association (IEA/HFES '00)*, vol. 44, pp. 286–289, Human Factors and Ergonomics Society, Santa Monica, CA, USA, August 2000.

[18] C. Forlines, B. Schmidt-Nielsen, B. Raj, K. Wittenburg, and P. Wolf, "A comparison between spoken queries and menu-based interfaces for in-car digital music selection," in *Proceedings of the International Conference on Human-Computer Interaction (INTERACT '05)*, pp. 536–549, Rome, Italy, 2005.

[19] L. Garay-Vega, A. K. Pradhan, G. Weinberg et al., "Evaluation of different speech and touch interfaces to in-vehicle music retrieval systems," *Accident Analysis and Prevention*, vol. 42, no. 3, pp. 913–920, 2010.

[20] U. Gärtner, W. König, and T. Wittig, "Evaluation of manual vs. speech input when using a driver information system in real traffic," in *Proceedings of the Driving Assessment 2001: The First International Driving Symposium on Human Factors in Driving Assessment, Training and Vehicle Design*, Aspen, Colo, USA, 2001.

[21] K. Itoh, Y. Miki, N. Yoshitsugu, N. Kubo, and S. Mashimo, "Evaluation of a voice-activated system using a driving simulator," SAE World Congress & Exhibition, SAE Tech 2004-01-0232, Society of Automotive Engineers, Warrendale, Pa, USA, 2004.

[22] J. Maciej and M. Vollrath, "Comparison of manual vs. speech-based interaction with in-vehicle information systems," *Accident Analysis and Prevention*, vol. 41, no. 5, pp. 924–930, 2009.

[23] M. C. McCallum, J. L. Campbell, J. B. Richman, J. L. Brown, and E. Wiese, "Speech recognition and in-vehicle telematics devices: potential reductions in driver distraction," *International Journal of Speech Technology*, vol. 7, no. 1, pp. 25–33, 2004.

[24] T. A. Ranney, J. L. Harbluk, and Y. I. Noy, "Effects of voice technology on test track driving performance: implications for driver distraction," *Human Factors*, vol. 47, no. 2, pp. 439–454, 2005.

[25] J. Shutko, K. Mayer, E. Laansoo, and L. Tijerina, "Driver workload effects of cell phone, music player, and text messaging tasks with the Ford SYNC voice interface versus handheld visual-manual interfaces," SAE World Congress & Exhibition, SAE Tech 2009-01-0786, Society of Automotive Engineers, Warrendale, Pa, USA, 2009.

[26] O. Tsimhoni, D. Smith, and P. Green, "Destination entry while driving: speech recognition versus a touch-screen keyboard," Tech. Rep. UMTRI-2001-24, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 2002.

[27] J. Villing, C. Holtelius, S. Larsson, A. Lindstrom, A. Seward, and N. Aberg, "Interruption, resumption and domain switching in in-vehicle dialogue," in *Proceedings of the 6th International Conference on Natural Language Processing*, pp. 488–499, Gothenburg, Sweden, 2008.

[28] V. E.-W. Lo, P. A. Green, and A. Franzblau, "Where do people drive? Navigation system use by typical drivers and auto experts," *Journal of Navigation*, vol. 64, no. 2, pp. 357–373, 2011.

[29] U. Winter, T. J. Grost, and O. Tsimhoni, "Language pattern analysis for automotive natural language speech applications," in *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (ACM '10)*, pp. 34–41, Pittsburgh, Pa, USA, November 2010.

[30] K. Takeda, J. H. L. Hensen, P. Boyraz, L. Malta, C. Miyajima, and H. Abut, "International large-scale vehicle corpora for research on driver behavior on the road," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, pp. 1609–1623, 2011.

[31] A. Barón and P. A. Green, "Safety and usability of speech interfaces for in-vehicle tasks while driving: a brief literature review," Tech. Rep. UMTRI-2006-5, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 2006.

[32] B. Faerber and G. Meier-Arendt, "Speech control systems for handling of route guidance, radio and telephone in cars: results of a field experiment," in *Vision in Vehicle—VII*, A. G. Gale, Ed., pp. 507–515, Elsevier, Amsterdam, The Netherlands, 1999.

[33] A. Kun, T. Paek, and Z. Medenica, "The effect of speech interface accuracy on driving performance," in *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech '07)*, pp. 1326–1329, Antwerp, Belgium, August 2007.

[34] A. W. Gellatly and T. A. Dingus, "Speech recognition and automotive applications: using speech to perform in-vehicle tasks," in *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*, pp. 1247–1251, Santa Monica, Calif, USA, October 1998.

[35] R. M. Schumacher, M. L. Hardzinski, and A. L. Schwartz, "Increasing the usability of interactive voice response systems: research and guidelines for phone-based interfaces," *Human Factors*, vol. 37, no. 2, pp. 251–264, 1995.

[36] *Intuity Conversant Voice Information System Version 5. 0 Application Design Handbook*, AT&T Product Documentation Development, Denver, Colo, USA, 1994.

[37] L. J. Najjar, J. J. Ockeman, and J. C. Thompson, "User interface design guidelines for speech recognition applications," presented at IEEE VARIS 98 Workshop, Atlanta, Ga, USA, 1998, http://www.lawrence-najjar.com/papers/User_interface_design_guidelines_for_speech.html.

[38] Z. Hua and W. L. Ng, "Speech recognition interface design for in-vehicle system," in *Proceedings of the 2nd International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 29–33, ACM, Pittsburgh, Pa, USA, 2010.

[39] "Ergonomics—Assessment of Speech Communication," ISO Standard 9921, 2003.

[40] "Ergonomics—Construction and Application of Tests for Speech Technology," Tech. Rep. ISO/TR 19358, 2002.

[41] "Information Technology—Vocabulary—Part 29: Artificial Intelligence—Speech Recognition and Synthesis," ISO/IEC Standard 2382-29, 1999.

[42] "Acoustics—Audiometric Test Methods—Part 3: Speech Audiometry," ISO Standard 8253-3, 2012.

[43] "Ergonomics of human-system interaction—Usability methods supporting human-centered design," ISO Standard 16982, 2002.

[44] "Road vehicles—Ergonomic aspects of transport information and control systems—Specifications for in-vehicle auditory presentation," ISO Standard 15006, 2011.

[45] *Voice User Interface Principles and Guidelines (Draft)*, SAE Recommended Practice J2988, 2012.

[46] R. Hopper, *Telephone Conversation*, Indiana University Press, Bloomington, IN, USA, 1992.

[47] B. Balentine and D. P. Morgan, *How to Build A Speech Recognition Application*, Enterprise Integration Group, San Ramon, Calif, USA, 1999.

[48] M. H. Cohen, J. P. Giangola, and J. Balogh, *Voice Interface Design*, Pearson, Boston, Mass, USA, 2004.

[49] R. A. Harris, *Voice Interaction Design*, Morgan Kaufmann, San Francisco, Calif, USA, 2005.

[50] J. R. Lewis, *Practical Speech User Interface Design*, CRC Press, Boca Raton, Fla, USA, 2011.

[51] S. C. Levinson, *Pragmatics*, Cambridge University Press, New York, NY, USA, 1983.

[52] G. Skanztze, "Error detection in spoken dialogue systems," 2002, http://citeseer.ist.psu.edu/cache/papers/cs/26467/http:zSzzSzwww.ida.liu.sezSz~nlplabzSzgsltzSzpaperszSzGSkantze.pdf/error-detection-in-spoken.pdf.

[53] J. L. Austin, *How To Do Things With Words*, Harvard University Press, Cambridge, Mass, USA, 1962.

[54] A. Akmajian, R. A. Demers, A. K. Farmer, and R. M. Harnish, *Linguistics: An Introduction To Language and Communication*, MIT Press, Cambridge, Mass, USA, 5th edition, 2001.

[55] J. R. Searle, "A taxonomy of illocutionary," in *Language, Mind and Knowledge, Minnesota Studies in the Philosophy of Science*, K. Gunderson, Ed., vol. 7, pp. 344–369, 1975.

[56] H. P. Grice, "Logic and conversation," in *Syntax and Semantics 3: Speech Acts*, P. Coole and J. L. Morgan, Eds., pp. 41–58, Academic Press, New York, NY, USA, 1975.

[57] J. Véronis, "Error in natural language dialogue between man and machine," *International Journal of Man-Machine Studies*, vol. 35, no. 2, pp. 187–217, 1991.

[58] N. Chomsky, *Aspects of Theory of Syntax*, The MIT Press, Cambridge, Mass, USA, 1965.

[59] M. L. Bourguet, "Towards a taxonomy of error-handling strategies in recognition-based multi-modal human-computer interfaces," *Signal Processing*, vol. 86, no. 12, pp. 3625–3643, 2006.

[60] C.-M. Karat, C. Halverson, D. Horn, and J. Karat, "Patterns of entry and correction in large vocabulary continuous speech recognition systems," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 568–575, ACM, Pittsburgh, Pa, USA, May 1999.

[61] K. Larson and D. Mowatt, "Speech error correction: the story of the alternates list," *International Journal of Speech Technology*, vol. 6, no. 2, pp. 183–194, 2003.

[62] D. Litman, M. Swerts, and J. Hirschberg, "Characterizing and predicting corrections in spoken dialogue systems," *Computational Linguistics*, vol. 32, no. 3, pp. 417–438, 2006.

[63] E.-W. Lo, S. M. Walls, and P. A. Green, "Simulation of iPod music selection by drivers: typical user task time and patterns for manual and speech interfaces," Tech. Rep. UMTRI-2007-9, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 2007.

[64] J. F. Kelley, "An empirical methodology for writing user-friendly natural language computer applications," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 193–196, ACM, Boston, Mass, USA, 1983.

[65] J. D. Gould, J. Conti, and T. Hovanyecz, "Composing letters with a simulated listening typewriter," *Communications of the ACM*, vol. 26, no. 4, pp. 295–308, 1983.

[66] P. Green and L. Wei-Hass, "The Wizard of Oz: a tool for repaid development of user interfaces," Tech. Rep. UMTRI-1985-27, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 1985.

[67] A. K. Sinah, S. R. Klemmer, J. Chen, J. A. Landay, and C. Chen, "Suede: iterative, informal prototyping for speech interfaces," in *Proceedings of the CHI 2001 Proceedings*, Association for Computing Machinery, New York, NY, USA, 2011.

[68] L. Dybkjær, N. O. Bernsen, and W. Minker, "Evaluation and usability of multimodal spoken language dialogue systems," *Speech Communication*, vol. 43, no. 1-2, pp. 33–54, 2004.

[69] M. Walker, C. Kamm, and D. Litman, "Towards developing general models of usability with PARADISE," *Natural Language Engineering*, vol. 6, pp. 363–377, 2000.

[70] M. Hajdinjak and F. Mihelič, "The PARADISE evaluation framework: issues and findings," *Computational Linguistics*, vol. 32, no. 2, pp. 263–272, 2006.

[71] J. Schweitzer and P. A. Green, "Task acceptability and workload of driving urban roads, highways, and expressway: ratings from video clips," Tech. Rep. UMTRI-2006-6, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 2007.

[72] P. Green, B. T.-W. Lin, J. Schweitzer, H. Ho, and K. Stone, "Evaluation of a method to estimate driving workload in real time: watching clips versus simulated driving," Tech. Rep. UMTRI-2011-29, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 2011.

[73] P. Green, "Using standards to improve the replicability and applicability of driver interface research," in *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular (AutomotiveUI '12)*, Portsmouth, UK.

[74] M. R. Savino, *Standardized names and definitions for driving performance measures [Ph.D. thesis]*, Department of Mechanical Engineering, Tufts University, Medford, Ore, USA, 2009.

[75] *Operational Definitions of Driving Performance Measures and Statistics (Draft)*, SAE Recommended Practice J2944, 2012.

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2013, Article ID 705086, 12 pages
http://dx.doi.org/10.1155/2013/705086

Research Article

Evaluation of a Navigation Radio Using the Think-Aloud Method

Paul A. Green[1] and Jin-Seop Park[2]

[1] *Driver Interface Group, University of Michigan Transportation Research Institute, 2901 Baxter Road, Ann Arbor, MI 48109-2150, USA*
[2] *Advanced Vehicle Safety and Dynamics Research Office, Korea Automobile Testing and Research Institute, 200 Samjon-ro, Songsan-myeon, Hwaseong-si, Gyeonggi-do 445-871, Republic of Korea*

Correspondence should be addressed to Paul A. Green; pagreen@umich.edu

Received 10 October 2012; Accepted 26 December 2012

Academic Editor: Motoyuki Akamatsu

Copyright © 2013 P. A. Green and J.-S. Park. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this experiment, 13 licensed drivers performed 20 tasks with a prototype navigation radio. Subjects completed such tasks as entering a street address, selecting a preset radio station, and tuning to an XM station while "thinking aloud" to identify problems with operating the prototype interface. Overall, subjects identified 64 unique problems with the interface; 17 specific problems were encountered by more than half of the subjects. Problems are related to inconsistent music interfaces, limitations to destination entry methods, icons that were not understood, the lack of functional grouping, and similar looking buttons and displays, among others. An important project focus was getting the findings to the developers quickly. Having a scribe to code interactions in real time helped as well as directed observations of test sessions by representatives of the developers. Other researchers are encouraged to use this method to examine automotive interfaces as a complement to traditional usability testing.

1. Introduction

People want products that are easy to use, and that is partic-ularly true of motor vehicles. Numerous methods have been developed to assess the ease of use of driver interfaces, both traditionally, and more recently from the human-computer interaction literature [1–3]. The three most prominent methods are (1) usability testing [4–8], (2) expert reviews [9–11], and (3) the think-aloud method [12–15]. Methods vary in terms of their value for formative evaluation (while development is in progress) and summative evaluation (at the end of development). See [16] for an extensive overview of how various methods are conducted and where they should be applied.

Usability testing is the gold standard of usability test methods, as it involves real users performing real tasks, though often in a laboratory setting, and can be part of either formative or summative testing. The purpose is to determine task completion times and errors. Generally, usability testing occurs in the latter stages of design, when a fully functioning interface is available. Usability tests are time-consuming to plan and analyze and can be costly.

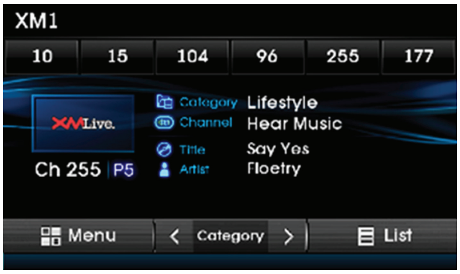
Consequently, there has been considerable interest in predicting user performance, in particular task time [17–22]. Task times for experienced users can be predicted in a fraction of the time to plan, conduct, and analyze a usability test. If the method used by subjects to perform a task is known, the predictions should be as accurate as the usability test data [23].

Expert reviews can be an efficient alternative to usability testing, especially early in design, though they may be used for summative testing as well. In an expert review, each step of each task is examined to determine how the interface should be designed according to established usability heuristics and guidelines. Expert reviews are often criticized as being "just someone's opinion." Therefore, reviewers should be professionally certified in human factors or usability. (See http://www.bcpe.org/).

In the think-aloud method, users describe their logic as they try to use an interface. For example, a subject might say, "I selected the city name but cannot figure out how to get to the next step," or "Sometimes, there is an OK button in the lower right corner, but there is not one here.

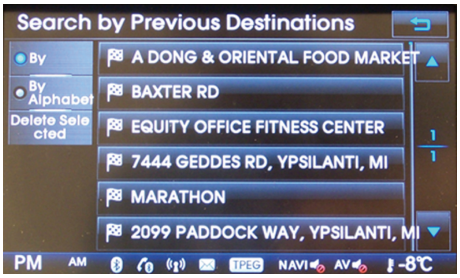
Figure 1: Most common screen.

Figure 2: One of two destination screens.



Category Lifestyle
Channel Hear Music
Title Say Yes
Artist Floetry

Figure 3: XM radio screen.



A DONG & ORIENTAL FOOD MARKET
BAXTER RD
EQUITY OFFICE FITNESS CENTER
7444 GEDDES RD, YPSILANTI, MI
MARATHON
2099 PADDOCK WAY, YPSILANTI, MI

Figure 4: List of previous destinations screen.

I am stuck and frustrated." The think-aloud method helps evaluators identify what is confusing or misleading and how those problems can be resolved. If subjects fall silent, then the experimenter prompts them to speak but needs to do without interfering with the subject's thinking process or influencing them. When to prompt and what to say is much more difficult to do than it may seem. In fact, considerable experimenter training and practice are required. See [16] for a discussion.

Think-aloud evaluations are most useful during the early stages of design while the design is still being formulated, identifying problems users experience more readily than other methods. Unfortunately, as is described later, data reduction in think-aloud evaluations is very time-consuming.

Recently, the authors conducted an evaluation of an early prototype of the Mobis Generation 3 navigation radio for Hyundai-Kia vehicles [24]. The complete experiment included four parts (1) a think-aloud evaluation from a human factors expert, (2) a think-aloud evaluation involving 13 ordinary drivers, (3) a follow-up survey of those drivers primarily concerning their understanding of icons, and (4) estimates of task times per SAE Recommended Practice J2365 [20]. Because this experiment was conducted during the early stage of interface design, and the interface designers needed to know what problems users would encounter, the focus of the experiment and this paper concerns only part 2, the think-aloud evaluation by ordinary drivers.

There are many other ways this data could have been collected. For example, questions concerning what subjects did and what subjects did and why, could have been asked retrospectively. In selecting methods to utilize, the authors considered the specifics of the request for quote from the sponsor, verifying conformance to accepted industry practice (e.g., SAE J2365), what information was believed to be most useful to the sponsor, the experience of the research team, the funding for the project, the schedule, and other factors. There was extreme pressure to complete this project very quickly to meet the production schedule set by Hyundai-Kia. Therefore, considerable thought was given as to how to complete this project quickly, which meant that less time was spent on certain activities than is ideal, and methods to accelerate data collection and analysis were explored.

2. Method

2.1. Navigation Device Examined. The device examined was an early working prototype of a Mobis generation 3 navigation radio. As shown in Figure 1, the navigation radio consisted of an LCD display surrounded by 10 hard buttons (e.g., select satellite radio, seek), two CD related buttons, a volume knob, and a tuning knob. These hard buttons as well as soft buttons on the touch screen allowed access to hundreds of screens. Figures 2, 3, and 4 show example screens.

2.2. Test Facility. The experiment was performed using the third generation UMTRI driving simulator while "parked." The navigation system was mounted into the center stack of the simulator cab. To enable signal reception and use of the GPS and XM functions, an antenna was installed, connecting the simulator lab with an outside room. Figures 5 and 6 show a hypothetical subject being recorded, the equipment used, and the recorded image from an actual subject. Although the cameras were in plain sight, subjects ignored them, in part because the camera in front of them was small.

2.3. Sequence of Tasks in the Experiment. In each session, one session per subject, subjects (1) completed biographical and consent forms and had their vision checked, (2) practiced the

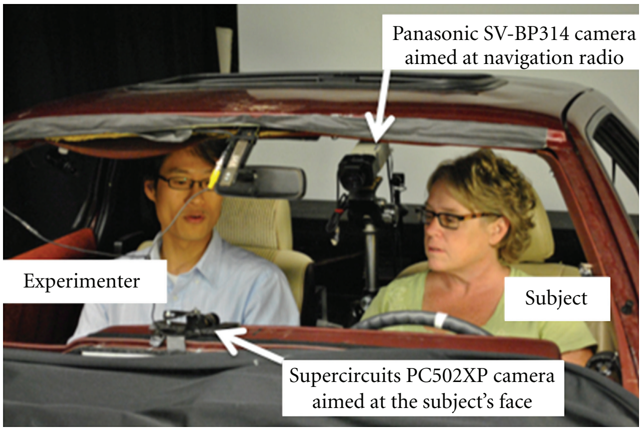


Figure 5: Experimenter and a hypothetical subject interacting. Note: not shown are the Panasonic WJ-420 quad splitter, the Sharp LC10-A3U-B 10 in LCD monitor, the Panasonic DMR-EZ47V DVD recorder, the AudioTechnica microphone, or the Mackie DX8 digital audio mixer.

Figure 6: Sample screen grab from a typical subject.

think-aloud method, (3) completed 20 test tasks in a fixed order while thinking aloud, and (4) completed a 19-page survey requested by the sponsor. Practice involved counting the number of chairs in the place where they lived, going room by room. As intended, the subjects did not simply list how many chairs were in each room, but said something about the kind of chairs present, where the chairs were located in each room, or other information. If they did not provide some of those details, then a question was asked to encourage them. Often, the process of recalling chairs was a virtual journey ("The front door takes me into the living room. In that room… Next to that room, around the table …"). The practice was quick and helped subjects understand what was meant by "think aloud."

Some 20 tasks were examined (Table 1). These tasks were selected because of their importance, frequency of occurrence, and to provide data on a variety of entry types. Only manual entry was allowed. For ease of administration and analysis, the order of tasks was fixed.

Subjects were not given any documentation or instruction as to how the interface functioned, as the interface design was intended to be intuitive.

The second author served as the experimenter for the main experiment. He had reviewed the literature on think-aloud studies and also served as the experimenter for pilot testing of the first author, which led to discussions of when to prompt and what to say. The second author also observed and provided feedback on test sessions of the first few subjects after each session was complete. Ideally, more time would have been desired for training, but that was not feasible given the sponsor's product development schedule.

During the think-aloud test, the experimenter, seated in the front passenger seat, presented a sheet of paper, one per task, describing the task to perform and the data to use, and asked the subject to think aloud while doing the task. When subjects fell silent for an extended period of time (typically 30 s, but sometimes longer), the experimenter prompted them. ("What are you looking for? What are you expecting to see? Are you confused?") Generally, few prompts were needed. There were no specific rules about which prompt to use when. Although these prompts seem leading, it was often apparent from what subjects said and their actions what that the subject state was consistent with these prompts. For example, if a subject repeatedly switched between screens but did not select anything else, then "What are you looking for" was an appropriate prompt.

If a subject was not making any progress after three to five minutes on a task (depending on the time remaining), they were given a hint. Hints identified what to do next (e.g., press this button), without any explanation of why an action was appropriate. If subjects continued to struggle, they were told to stop and move on to the next task, as the intent was to reveal as many problems as possible in the time available.

A scribe (a very fast typist) sat outside the simulator cab and attempted to record verbatim what the experimenter and subject said during the experiment. The only and very general instructions to the scribe were to record everything said and to use the video and audio recordings to fill in any gaps. In part, a verbatim transcription was feasible because the experimenter and subject were not talking continuously, so the scribe was able to catch up during pauses. Anything that the scribe missed (e.g., if the subject spoke quickly) was filled in immediately after each test session (in the 30 minutes or so before the next subject) from the audio and video recordings. Having a complete transcription essentially immediately after each session was completed shortened the time to reduce the data. Furthermore, after the fact review of the audio recordings, collected using a system that was hastily assembled and not optimized for recording quality, were sometime inaudible, especially when subjects mumbled. Although requiring a scribe in addition to an experimenter made the experiment more difficult to schedule, given when this experiment was conducted, a scribe was always available.

2.4. Subjects. Thirteen licensed drivers volunteered to serve as subjects: 6 younger people (ages 19–26, 3 men and 3 women) and 7 older people (ages 65–83, 3 men and 4 women). They were recruited via an advertisement on Craigslist. All older subjects were retired. Five of the younger subjects were students. Younger subjects were the most likely users of the navigation radio, especially of the audio functions. Older subjects were those most likely to be challenged by the interface and would be the first to encounter problems

Table 1: Task list.

Category	#	Task	Data to use
Navigation	1	Enter destination via street address method	100 North 5th Avenue; Ann Arbor, MI
	2	Enter waypoint via POI (Point Of Interest) method	Nearest McDonalds from current position
	3	Enter destination via POI method	Nearest gas station from current position
	4	Cancel the route	Current route
	5	Enter destination via POI method	Detroit Metro Airport
	6	Change the map scale	Zoom in
	7	Enter destination via street address method	1600 Pennsylvania Avenue; Washington, DC
	8	Enter destination via intersection method	US-23 and I-94
	9	Enter destination via POI method	Cobo Hall in Detroit, MI
	10	Enter destination via intersection method	Broadway St. and Plymouth Rd.
	11	Store address	Subject’s home address
CD	12	Insert CD and select track	CD provided, third track
FM/AM radio	13	Tune radio	FM 91.7
	14	Adjust volume	To a comfortable level
	15	Set preset	FM 91.7 as the Preset 1
	16	Adjust volume	To a comfortable level
Satellite radio	17	Tune XM radio	Elvis Station
	18	Turn off XM radio	Turn off
Clock	19	Change and set time	11:45 pm
iPod	20	Connect iPod and select song	“Wonderful Tonight”

this study was to reveal. All subjects had corrected visual acuity adequate to drive. They drove a mean of 8,000 miles per year, somewhat less than is typical in the United States.

Other than being a licensed driver, in good health, in specific age categories, a native English speaker, and experience with XM/Sirius radio, there were no other requirements to participate. Thus, there was no control over experience with technology in this experiment, a limitation not included so subjects could be recruited in the time frame available. Experience with relevant technology was mixed. Five subjects owned an iPod, and one owned another brand of MP3 player. Only two of the subjects had vehicles with XM/Sirius radios. Five subjects owned GPS systems: 3 TomToms and 2 Garmins.

Subjects were paid $40 for their time if they completed the experiment in two hours. Subjects who took longer (some of the older subjects) were paid an extra $10.

The sample, seemingly small, was more than adequate for identifying problems, of which 64 were identified. “A problem was defined as a situation where a task took too long, subjects struggled to make progress, or they otherwise expressed doubt (“I am not sure what button to press”), confusion, irritation (“I would like to shoot the engineer that designed this”), or other undesired feelings” ([24] page 48). Problems invariably involved deviations from the intended sequence of steps to complete the task expeditiously. Research shows that after about six subjects or so, the number of new problems found with each additional subject is small, with the specific number varying the problem severity and other factors [25, 26]. Specifically, the six subject value comes from assuming that each subject has about a one in three chance of discovering a problem, that problems are independent, and the goal is to discover 90% of the problems. Furthermore, this initial analysis assumes all problems are of a similar severity, and one may have different goals for different level of severity. (See [27] for the most recent of a long series of papers, reports, and now a book chapter on sample size.)

Although one can quibble on the specifics of the calculations, the surprise to many nonhuman factors, nonusability professionals is that most of the problems can be found with just a few more than a handful of properly selected subjects. The data from this experiment confirmed that conclusion (Figure 7), with most problems being found by the first few subjects. Testing more subjects would provide better statistical evidence for the frequency of occurrence of each problem and would identify more problems, but not many. Most importantly, testing more subjects would delay producing a complete report informing the designers of interface problems to be corrected. In this case, boosting the confidence of the sponsor was partly why more than six subjects or so were tested. Further, in these instances, when deciding which aspects of an interface to modify, the percentage of subjects who encounter a problem may be secondary. Rather, if just one subject encounters a problem, and the problem seems reasonable, then changes to eliminate or reduce the impact of that problem should be considered.

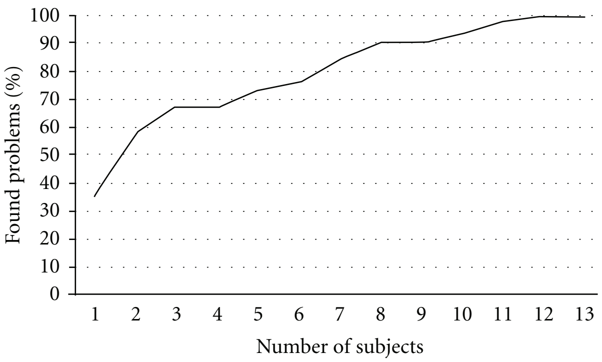


Figure 7: Cumulative usability problems found as a function of number of subjects. Note: this figure determines the % of problems (out of 64 total) found utilizing data from subjects in the order they were tested. So, combining the data of subjects 1 and 2, 38/64 = 59% of the problems were found.

3. Results

3.1. Data Reduction for Problem Identification. Data reduction consists of (1) listening to each session to verify each transcription, correcting them as needed, (2) identifying each problem that users experienced, and (3) identifying the frequency, severity, and persistence of each problem. A problem was indicated when a task took too long, subjects struggled to make progress, or they otherwise expressed doubt ("I am not sure what button to press"), confusion, or irritation ("I would like to shoot the engineer that designed this").

Subjects indicated problems in several ways. Indications included (1) questions ("What is this? Where is the button? How do I get the map?"), (2) statements of uncertainty ("I'm not sure if I can have the map when I click on this button. It does not allow me to save this radio frequency. I am wondering why it's not accepting this."), and (3) exclamations with filler words ("Oh, man…," "Umm, …"). About 13% of the problems for younger subjects were associated with questions, whereas for older subjects, questions were linked to a third of the problems. In contrast, about 81% of the problems for younger subjects were associated with statements of uncertainty, versus 61% of the problems for older subjects. There was no difference in the use of filler words (about 7% for both age groups). Specific examples of how particular problems identified appear later in this paper.

There were instances during the experiment where the subject was silent for an extended period of time, where there were no probes from the experimenter, and where it was uncertain from the transcript and video recording what the subject was thinking. This often occurs with novice experimenters as they focus on observing what subjects do. One solution would be a timing device prompted the experiment (unobtrusively to the subject) to probe the subject to think aloud.

Frequency is the number of times the problem occurred, usually across subjects, but sometimes within subject groups. Persistence can be the number of times a problem reoccurred within each subject.

There were three primary severity categories—critical ("showstopper"), major, and minor (often cosmetic). A critical problem prevents subjects from completing a task, such as not finding a power button or an enter key. A major problem substantially delays the subject but can be overcome. A minor problem has minimal impact on performance, but a change is nonetheless desired, such as making a label a different color or choosing a different font. For this study, task time was used to determine severity. Task times greater than 300 s (5 minutes) were critical. Greater than 30 s but less than or equal to 300 s was major. Minor was 30 s or less.

In the literature are a number of formal methods to analyze verbal protocols some of which the authors were unaware of at the time this experiment was conducted (e.g., [28]). However, for the purpose of this applied analysis, given the degree to which the subjects were expressive the experience of the experimenter, and the time available, a custom categorization scheme seemed more appropriate. In brief, problems were categorized as device domain or subject domain problems. Device domain problems were (1) visual or auditory interface related (layout, label, text, sound, and action time), (2) logic and organization (controls, search, system, and information architecture), or (3) nonusability software issues (stability, database, and response time). Subject problems included data knowledge, procedure knowledge, and preference. In this case, the scheme was to aid Korean designers, many of who were not fluent in English, and had no human factors background.

Finally, given focus on what the sponsor needed to fix, the limited time schedule, limited funding, no effort was made to examine the effect of subject differences (young versus old, those who had navigation systems versus those who did not, etc.) on the particular problems encountered. Although very interesting, they were secondary issues.

3.2. How Often Were Subjects Able to Complete the Tasks? Table 2 shows how well subjects did when they were given hints. One could interpret this data to suggest that the interface was relatively easy to use, but that was not the case. *Many subjects required hints to complete many of the tasks*. Had hints not been provided, the success rate would have been half the values shown, or typically less than 50%, quite poor. For perspective, subjects took 2 hours on average to complete these 20 tasks, which is about 6 minutes/task, a long time. Several of the tasks, such as setting a preset radio frequency, should only take a few seconds to complete.

3.3. How Often Did Problems Occur and How Severe Were They? Table 3 lists the 64 problems, a rather large number, from the most frequent to the least frequent. About two-thirds of all problems were experienced by at least 2 subjects. Among those, 16 problems (25 percent of the total) were encountered by more than half of all subjects. In the report summarizing this project [24] the frequency of problems was reported many ways. For the table that follows, they are reported at the number of subject (out of 13) experiencing a problem because this was the format most readily understood by the sponsor and its interface designers. Furthermore, the experiment was deliberately designed to minimize repetition

Table 2: Successful completion with hints.

#	Task	Success #	Success %	# Not do
1	Set destination to 100 north 5th avenue, Ann Arbor, MI.	9	69	
2	Find the nearest McDonalds and set it as a waypoint.	8	62	
3	Find the nearest gas station from the current position.	11	85	
4	Cancel the route that you have set up.	10	77	
5	Enter and set a destination to the main Detroit Airport.	12	92	
6	Change map scale to zoom in slightly.	10	77	
7	Enter and set a destination to 1600 Pennsylvania Ave, Washington DC.	10	77	
8	Set the destination to the intersection of US-23 and I-94.	1	8	
9	Set the destination to Cobo Hall in Detroit, MI.	9	100	4
10	Set the destination to intersection of Broadway St and Plymouth Rd, Ann Arbor.	9	100	4
11	Enter your home address and store it.	3	50	7
12	Insert the CD and select the third track.	13	100	
13	Adjust volume to a comfortable level.	13	100	
14	Tune FM radio to 91.7.	13	100	
15	Adjust volume to a comfortable level.	13	100	
16	Save the frequency as first preset.	6	46	
17	Tune XM radio to Elvis station.	10	77	
18	Turn off XM radio.	12	100	1
19	Change the time to 11:45 pm.	13	100	
20	Connect the iPod and select the song “Wonderful Tonight.”	12	92	

Note: under the heading success, # refers to the number of subjects out of 13 who successfully completed the task with hints. Thus, for task 1, 100 × 9/13 = 69%. The # Not Do refers to the number of subjects who were not asked to do up to four tasks to keep the experiment from taking too long. Thus, for task 9, only 9 of the 13 subjects performed the task, but all 9 of them (100%) complete the task.

of tasks and task elements, and thereby more widely explore the interface, which tend to limit repeated encounter with problems.

The linkage between what was observed and these problems can best be described by example. Following is a description of some of the most frequent and critical problems. Problems were identified by a combination of how long it took to complete each task and step (if they were completed at all), what subjects did, what subjects said, and although not described here, their facial expressions. The ultimate indication of a problem was when they said, “I give up.” Quite frankly, identifying when a problem had occurred was fairly obvious.

The most frequent problem (9) was that the system did not accept “DC” as a state name when subjects searched for an address in Washington, D.C (task 1). All subjects except one (92% of the sample) tried to type “DC.” They would get to the state field and type in “D.” Immediately, the C key would gray out because the system was expecting the subject to type “District.” “Why is the C key gray? I want to type it.” In contrast to most state names, the District of Columbia is invariably abbreviated as “DC,” but the software accepted only “District of Columbia.” Allowing the intelligent speller to accept both complete names and two-letter abbreviations would eliminate this problem, and speed the entry of other state names as well.

The system often froze (problem 10) while the experiment was in progress (for 11 of the 13 subjects, and for some subjects, multiple times). Given that a prototype was being evaluated, some problems were expected. The work around the freeze was to unplug the system, plug it back in again, and wait for the system to restart. This process took a minute or so to complete and was mildly annoying to subjects and experimenters. Although testing could have been conducted when the interface had fewer bugs that froze the interface, that testing would have occurred later in design when there was less time and fewer resources to correct problems identified.

For task 16, 10 of the 13 subjects could not figure out how to set the radio presets (push and hold the soft button), problem 42. When they reached the preset screen, they would say something such as “Where is the set button?” Often, they would push “autostore,” which reset all the presets. In fact, one subject did this three times, exclaiming, “Why did it do that, again!” when the desired frequency setting appeared on a different button than it was prior to pressing autostore. Subjects did not realize that the method used to preset a radio frequency that works for radios with mechanical buttons (push and hold), also worked here.

Some 9 of the 13 subjects had trouble understanding the label to change the keyboard from alphabetic mode (the default) to numeric mode (problem 2), a step required to

Table 3: List of problems by frequency for all 13 subjects.

Frequency	Problem #	Severity	Description
12	9	Major	State name cannot be abbreviated (DC, Washington DC)
11	10	Minor	System freeze
10	14	Minor	No way to enter highway intersection as destination
10	42	Major	No intuitive method to save radio frequency as preset
9	2	Major	Difficult to switch keyboard to number keys
9	20	Major	Could not search for an XM channel by name, so they browsed the numeric listing, which was slow
9	24	Critical	Typed building number in the street name (north/no)
9	35	Major	Searching for an iPod song by title was very time-consuming
8	4	Minor	Guidance voice quality was inadequate
8	6	Critical	No intuitive method to cancel route/destination on map screen
8	7	Major	Buttons are easily confused with displays
8	8	Major	Expected this interface to support multitouch, in particular 2-finger zooming
8	17	Minor	Correct address rejected (wrong house number range)
7	28	Major	Did not realize what the search button did (also used for enter)
7	29	Major	Incomplete and insufficient text information
7	37	Critical	No alphabetic input method for numeric street name (Fifth Ave)
6	3	Minor	Negative response to autosuggest keyboard (autocomplete function)
6	13	Major	Did not realize there were multiple pages for destination screen
6	32	Major	Confusing intersection search guidance ("X", "space")
6	36	Major	No intuitive method to store the entered address/home address
6	45	Major	Did not understand icons on the map window (Z, compass, flag, current position, scale)
5	23	Major	Typed building number in the street name
5	26	Major	Did not understand the acronym POI
4	5	Major	Did not understand POI symbol/function on the map
4	16	Minor	Search displayed multiple entries for the same destination (3 Ann Arbor)
4	22	Major	Difficult to switch to alphabet keys
4	25	Critical	No intuitive method to search for a waypoint and add it as a waypoint on the map screen.
3	11	Minor	Too slow/fast system response time
3	12	Minor	Address/text information was not legible
3	27	Major	Did not recognize scale button on the map screen
3	30	Major	No way to back up (e.g., main mode screen, simulation mode, and local POI)
3	33	Minor	No intuitive method to bring up CD control after inserting CD
3	34	Major	Did not understand label SAT
3	38	Major	Did not understand label Zagat
3	39	Major	Did not recognize back button
3	40	Minor	Unable to get to CD mode after inserting CD
3	41	Major	Did not understand Autostore
3	44	Major	Did not understand icons on the status bar
3	57	Major	No intuitive graphic to select the searched result on the list (city/street)
2	46	Minor	Scale window disappears too quickly on the screen
2	48	Major	Narrow search coverage by POI Name (no support for McDonalds)
2	50	Major	Did not recognize the delete button on keyboard
2	64	Major	Did not understand label Map key
1	1	Minor	State name is highlighted and shown on Search by Address screen.
1	15	Minor	Confusing search result (Ann Arbor/Ann Arbor Twp)

Table 3: Continued.

Frequency	Problem #	Severity	Description
1	18	Minor	Misunderstood seek/track button label
1	19	Minor	Did not notice volume label
1	21	Major	No intuitive distinction between Route and Dest key
1	31	Minor	Wrong screen instruction (“search by POI” in intersection)
1	43	Minor	Did not understand label Dest
1	47	Minor	Select button location was inconsistent
1	49	Minor	Did not understand “Alternative” menu
1	51	Minor	Misunderstood Asian, Korean, and Chinese on POI list (as language)
1	52	Minor	No way to control volume on the screen
1	53	Minor	Did not understand preset scan
1	54	Minor	Misunderstood ATM and Bank on POI category list
1	55	Minor	Wrong unit information/temperature
1	56	Major	Did not understand play/stop icon
1	58	Major	Did not understand numbers next to the searched street names
1	59	Major	Not able to set clock in expected manner
1	60	Major	No intuitive layout for house number input
1	61	Minor	Did not understand orange line on the map
1	62	Minor	No autosave feature in entering city
1	63	Minor	Confusing categorization on POI list (gas station-travel/automotive)

Note: Problem #: The 64 problems were numbered 1–64.

complete task 1. They did not realize there was a key on the alphabetic keyboard screen to change modes, so they would search for other screens using the back key, proclaiming something like “where is the number screen?” At other times, they could say, “I do not know where it is, so I will try everything,” and they selected each key on the screen one by one to learn what each did. Some subjects who were methodical in their efforts found the alphabetic keyboard mode key in this manner. The source of the problem was that the mode keys/buttons looked like other buttons on the keyboard, and there was no spacing or graphics to group them apart. There were other buttons for which functional grouping, indicated by spacing and graphics, would have helped as well.

In fact, subjects had numerous problems with soft buttons. First, there was no graphic distinction between buttons and displays, so that when subjects were not sure what to do, they pressed everything, including displays. Providing buttons with a drop shadow or other unique graphical characteristic as well as auditory feedback (a click sound) when buttons are operated should reduce confusion. When driving, drivers should not be looking at the display. Many interfaces have a beep to confirm that a switch has been pressed, but that beep is often the same as the beep for an error, which confuses drivers.

Also noteworthy were two problems related to street address entry. Many American street names have a direction as part of the name (e.g., North Main) and subjects were uncertain if they should enter the street name as North Main, Main and select North, or abbreviate North as N or NO. Only one option was provided, so only 9 of the 13 subjects were able to complete entry of a street name containing a direction, even with hints. Often they tried using exactly the same and steps multiple times, saying something such as, “I think this is the way to do it, but I must have not done it quite right.” They did not realize the system would not accept the data as they entered it.

Subjects also struggled with street names that contained numbers (e.g., 5th street) as only one method of entry was supported, even though subjects may choose to enter those streets as numbers or alphabetically (Fifth). Numbered streets are quite common in the United States.

Related to this were problems associated with subjects not realizing what had been set or was a default. This was particularly true with setting the state. Keep in mind that many cities are located on rivers, because rivers provide both water and transportation. However, rivers also serve as a geographical boundary, so going to a nearby place may require changing the state.

Problems in search for songs and radio stations were common. In part, this was because the interface for each mode (AM/FM, XM, etc.) was unique. The criteria on which one could search was unique to the mode, and most importantly, so was the organization. What could be saved or preset varied with the mode, including what saves or presets were named. As a consequence, subjects needed to browse through pages of screens to find an XM radio station or song on a storage device, which was a very time-consuming process. Manually scrolling through a list of 100 or more items and reading them to find a desired item should not be done while driving, especially when the lists are in an order the subject cannot use to speed the search.

In summary, from most to least, problems included unclear labels (20 instances, such as that for the number key and the name autostore), problems with search (10, such as a lack of consistency in method names and methods available, especially in finding songs and XM stations), poor graphics (9, many icons were meaningless), disorganized system (9, information such as destination modes being split across 2 screens), illegible text (5, mostly text that was too small, especially on maps), poor layout (4, such as inconsistent location of the "ok" and "done" buttons), other organizational issues, unreliable software, and database errors (2 each, including missing addresses), and problems associated with unrecognizable sounds, slow system response (2 types), and disorganized controls (all 1 each).

3.4. Persistence. Using persistence to identify problems was less useful here than has been identified in the literature. Examples of the most persistent problems included not understanding the acronym POI, not understanding icons in the map window, system freeze, not knowing what the search button did, expecting multitouch to be supported, and not understanding the label Zagat. If anything, the persistence data reinforced the need for better icons and graphics, and potentially eliminating icons in some cases. As an example, there were two screens from which subjects could select a method to enter a destination. Often, they did not realize there was a second screen, so they were stuck. The icons were of no help. Had the icons been removed, leaving only text, all of the entry methods would be on one screen and user performance would likely have improved.

Although there are numerous navigation systems in use today, and many of them use icons, there are no standard icons for navigation functions in ISO Standard 2575 [29]. Although it may not be possible to develop well-understood icons for many of the navigation functions of interest, whatever is developed could be better than the current situation, where icons vary from system to system.

3.5. Combined Analyses. When presented with a huge list of problems, such as those in Table 3, the designers' immediate reaction is often to ignore the overwhelming user feedback. First, there is disbelief that subjects experienced all of the problems listed. Therefore, two representatives from the sponsor responsible for the interface design observed every subject, so they saw that the problems were real. Unfortunately, they were not native English speakers, so for the first few subjects they struggled to understand what was occurring. That was overcome by impromptu discussions with them between test sessions or at the end of the day to explain what was observed. Also provided were a few video outtakes for others not present. In a subsequent project, a secure web camera in the test room was provided so those not present could observe the experiment. In this case, the 13-hour time difference between the test facility (Ann Arbor, Michigan, USA) and the sponsor's main engineering center (Yongin-Shi, Gyunngi-Do, Korea) makes remote viewing inconvenient.

TABLE 4: Problem numbers listed summarized by frequency and severity.

	Severity		
Frequency	Critical ($t > 300$ s)	Major ($300 \geq t > 30$ s)	Minor ($30 \geq t$)
12		9	
11			10
10		42	14
9	24	2, 20, 35	
8	6	7, 8	4, 17
7	37	28, 29	
6		13, 32, 36, 45	3
5		23, 26	
4	25	5, 22	16
3		27, 30, 34, 38, 39, 41, 44, 57	11, 12, 33, 40
2		48, 50, 64	46
1		21, 56, 58, 59, 60	1, 15, 18, 19, 31, 43, 47, 49, 51, 52, 53, 54, 55, 61, 62, 63

Note: for the descriptions corresponding to the problem numbers listed here, see Table 3.

To help designers prioritize what they should do, tables were created using pairs of the dimensions of interest (frequency, severity, and persistence). Ideally, designers should consider those dimensions, the effort required to fix each problem, and the implications of fixing each problem on other problems, as well as other factors in making decisions about what to fix.

The frequency-severity table (Table 4) contains the most useful of the dimension combinations. The problems listed in the upper left area of the table (e.g., 9, 42, 24, 2, 20, and 35) are the most frequent and severe, and therefore, of the highest priority.

4. Conclusions and Discussion

There were many problems with this interface, some of which were expected early in design when this interface was examined. Subjects consistently needed hints to complete tasks, which was not indicative of an intuitive interface.

One might criticize the sponsor of the research for the existence of these problems. However, the more important point is that they supported qualified experts to examine their interface, identify problems, and suggest improvements. Keep in mind that it is not the researchers' role to make the changes desired, and what is changed represents a tradeoff between user impact, cost, schedule, and hardware and software limitations.

4.1. Lack of Style Guide. There were inconsistencies in the interface, for example, where the "done" or "ok" button was located (and how it was labeled). According to the sponsor's representatives, there was no style guide or other specific set of guidelines governing the interface, a situation that greatly increases the likelihood of problems due to inconsistency in the user interface. Creating a style guide, especially one based on research, is a major task, but well worth the effort. Most computer manufacturers have style guides to help ensure their interfaces are consistent (e.g., [30]).

4.2. Inconsistent Music Interfaces. The search methods available and what could be stored as presets varied with the media and are reflected in problems 16, 20, 28, 35, and others. These inconsistencies led to interfaces that were unique to each media, making interface navigation difficult. Admittedly, the underlying databases have different structures, but a more common format and more similar search features would have been beneficial, so that subjects would only need to learn one set of search methods that were consistently named, not a unique set of methods for each data set. Interestingly, this system (and most others) did not allow for aggregation of all favored music presets (AM/FM, XM, etc.) on a single screen.

4.3. Destination Entry Problems. There were numerous issues with destination entry, most due to limiting the ways in which information was to be entered and not making apparent what had already been set. These issues are reflected in problems 14, 16, 17, 37, and others.

4.4. Icons Were Not Understood. There was a desire to provide icons so the interface would be language independent and usable by a wider user group. That assumes the icons are understandable, which was not true here. Although better icons could be developed and included in ISO Standard 2575, how well they will be understood is uncertain. That suggests for some parts of the interface, icons may not be provided. In the case of selecting a destination entry method, all of the methods will fit on one screen instead of being distributed across two screens, making it easier for users to find the desired method. In other cases (e.g., maps), icons must be provided, as there is insufficient space for text (let alone icons plus text). However, in this instance, most icons were not understood (problems 44, 45, 64, etc.). In fact, in a very lengthy examination of the icons used in this interface, conducted after all tasks were completed, on average only 2 of the 13 subjects were able to correctly identify what the various icons meant when shown in context. Easy to understand map icons need to be developed.

4.5. Labels Were Not Understood. This included POI (problem 5), SAT (problem 34), Zagat (problem 38), and others. Each of the labels used should be considered and alternatives proposed. This needs to be done in conjunction with the effort to develop new icons as they are an alternative to text labels.

4.6. Lack of Functional Grouping. There were several instances where information on screens was not grouped by function, increasing the time for users to find particular information (and increasing errors as well). The best example of this is problem 2. Admittedly, space is extremely limited, but there were instances where spacing and graphics could have been utilized for this purpose.

4.7. Buttons and Displays Looked Alike. There were no common graphical elements to controls (primarily soft buttons) and displays (mostly icons), so that when subjects were lost, they pressed everything (problem 7). In such instances, having different auditory feedback for operation of a button and erroneous operation would have been helpful.

Many of these problems are not new (See [31]).

Beyond this specific interface, which does this experiment say about how think-aloud experiments should be conducted?

4.8. More Experimenter Training Needed. More time was needed to train the experimenter in the think-aloud method, in particular time spent on testing pilot subjects and reviewing video recordings of them. In this project, there were two days between when the interface actually worked and when testing had to begin to deliver results to meet the sponsor's production schedule, far too short. A minimum of two to three weeks is recommended. Training is particularly important for nonnative speakers of the language of subjects or those who do not have extensive experience in testing human subjects as experimenters. Typically, they do not prompt subjects enough, or more generally, just have problems in reading subjects, not knowing when to engage them. This need was reflected in silent periods, where neither the subject or the experimenter spoke, and often the experimenter just starred at the screen. In other situations, when designers without human factors/usability expertise conduct the testing, excessive leading of subjects is observed, typically involving telling subjects how to complete a task.

A list of probe questions and criteria for when those questions should be asked would be a useful addition to the training materials. Being repeatedly asked "what are you thinking" is annoying. Formal rules about when to intervene and what to do (just press this button, but not saying why) can be helpful.

Another idea is a device that would look for periods of silence and subject inaction. When those periods occurred for some time, the device would vibrate something on which the experimenter was sitting as a reminder to ask a probe question.

4.9. Real-Time Session Recording Helped. Secretaries or students who were fast typists sat in the room with the subject and the experimenter and served as the scribe, creating the session transcripts. Immediately after the session, when the session was fresh in their mind, the scribe checked their transcript against the audio and video recording. For this to occur, more than a few minutes is required between subjects. Nonetheless, generating transcripts in real time rather than after the fact from a video recording reduced the time to provide the results to the sponsor, which was important where a real evolving product with a rigid development schedule was concerned.

There was no evidence that the presence of a scribe was disruptive, which was a concern. For example, subjects did not look at the scribe or comment on what they were doing. Being nearby, the scribe could hear when the subject mumbled, which was often the case when people think aloud. The mumbling is often difficult to hear on recorded video, but is often the most informative part of an interaction. In this experiment, no software to support recording was used. Use of the Morae software to aid session recording [32] should be considered for the future.

There was also no time to acquire and set up a high-quality audio recording system. Thus, due to substandard audio quality, few segments could be used for outtakes. To partially overcome this problem, English subtitles of what the subjects and experimenter said were added. Time and cost permitting, even better for this audience would have been subtitles in both English and Korean.

4.10. Real-Time Observation Helped. The project team has always invited sponsors to watch test sessions directly. There are always concerns that additional observers will distract subjects, or they might do or say something that will interfere with a test protocol, requiring the data to be discarded (and additional time for replacement data to be collected). In on-road studies, seating for the subject, the experimenter, and equipment may leave only one unoccupied seat.

In this instance, there were two observers from the sponsor who sat quietly in the corner of the test room. Each night, they produced a summary that they sent to their designers, greatly shortening the time to provide feedback to them. The authors could not have responded as quickly. This feedback was an extremely important supplement to the written report and videos.

Had the hardware been available, providing a real time, web camera video of the experiment to the designers (in Korea) would have been useful. If the website URL is not advertised and a website is password protected, security should not be an issue. In this case, there were also issues with test sessions being a half-day out of synch with that of the designers and uncertainties about their ability to understand spoken English when it is disjointed and mumbled.

Surprisingly, the major challenge in getting the video to the users was not associated with the source, but getting the feed into the sponsor network because of security constraints. The solution is for the sponsor's employees to stay home when testing is expected, and watch it from there.

Keep in mind that if remote viewing is implemented, permission must be obtained from subjects on the consent form (and from the human subject board when the experiment is reviewed).

4.11. Data Reduction Was Slow. Creating a list of problems was a very slow and labor-intensive process, requiring the second author to read the transcripts and watch each session many, many times. Had there been time, the classification of problems would have benefited from a more structured approach, either considering them as problems related to goals and methods or to various error types (e.g., [33, 34]). A more structured approach could have shortened the time for data reduction, but only by a small amount. (See [16, 35–37].) If anything, a major weakness of the think-aloud method is that it takes so long to reduce the data, a concern when product development needs to be rapid.

4.12. More Information Identifying Solutions Was Needed. As the focus of the project was on identifying problems, less effort was given to identifying solutions. Had the time and resources been added, having a table that identified the solution to each problem (if there was a solution) could be useful.

The tone of these last comments may suggest the think-aloud method is seriously flawed. Quite to the contrary, the method provided extensive and detailed insights as to why users struggled to use the prototype navigation radio and how it could be improved. Wider use for automotive interface evaluation is strongly encouraged. Much of the information provided in this paper either could not have been obtained or would be difficult to obtain using other methods.

The ultimate test of this experiment was how the results were used. After being presented to Mobis, the sponsor, there was a follow-up presentation for Hyundai-Kia. The result was numerous requests from Hyundai-Kia to Mobis to modify the prototype interface to enhance its usability based on the experimental results. This was not a project that led to a report that just sat on a shelf. Furthermore, Hyundai-Kia has funded follow-on research now in progress.

Conflict of Interests

The authors declare that they have no conflict of interests.

Acknowledgment

This research was supported by a contract from Mobis to the University of Michigan.

References

[1] J. Nielsen and R. Mack, *Usability Inspection Methods*, John Wiley & Sons, New York, NY, USA, 1994.

[2] S. Krug, *Rocket Surgery Made Easy: The Do-It-Yourself Guide to Finding and Fixing Usability Problems*, New Riders, Berkeley, Calif, USA, 2010.

[3] J. Rubin, D. Chisnell, and J. Spool, *Handbook of Usability Testing*, John Wiley & Sons, New York, NY, USA, 2008.

[4] R. Jeffries, J. R. Miller, C. Wharton, and K. Uyeda, "User interface evaluation in the real world: a comparison of four techniques," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 119–124, 1991.

[5] R. A. Virzi, J. F. Sorce, and L. B. Herbert, "Comparison of three usability evaluation methods: Heuristic, think-aloud, and performance testing," in *Proceedings of the 37th Annual Meeting the Human Factors and Ergonomics Society*, pp. 309–313, October 1993.

[6] R. Molich, M. R. Ede, K. Kaasgaard, and B. Karyukin, "Comparative usability evaluation," *Behaviour and Information Technology*, vol. 23, no. 1, pp. 65–74, 2004.

[7] K. Hornbæk, "Current practice in measuring usability: challenges to usability studies and research," *International Journal of Human Computer Studies*, vol. 64, no. 2, pp. 79–102, 2006.

[8] M. W. M. Jaspers, "A comparison of usability methods for testing interactive health technologies: Methodological aspects and empirical evidence," *International Journal of Medical Informatics*, vol. 78, no. 5, pp. 340–353, 2009.

[9] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '90)*, Association for Computing Machinery, New York, NY, USA, 1990.

[10] A. M. Lund, "Expert ratings of usability maxims," *Ergonomics in Design*, vol. 5, no. 7, pp. 15–20, 1997.

[11] E. Hvannberg, E. Law, and M. Lárusdóttir M, "Heuristic evaluation: comparing ways of finding and reporting usability problems," *Interacting with Computers*, vol. 19, no. 2, pp. 225–240, 2007.

[12] C. H. Lewis and Using the, """Thinking Aloud" method in cognitive interface design," Technical Report IBM RC-9265, IBM Watson Research Center, Yorktown Heights, NY, USA, 1982.

[13] P. C. Wright and A. F. Monk, "The use of the Think-Aloud evaluation methods in design," *ACM SIGCHI Bulletin*, vol. 23, no. 1, pp. 55–57, 1991.

[14] E. Krahmer and N. Ummelen, "Thinking about thinking aloud: a comparison of two verbal protocols for usability testing," *IEEE Transactions on Professional Communication*, vol. 47, no. 2, pp. 105–117, 2004.

[15] M. Norgaard and K. Hornbaek, "What do usability evaluators do in practice? An exploratory study of Think-Aloud testing," in *Proceedings of the 6th conference on Designing Interactive systems (DIS '06)*, pp. 209–218, 2006.

[16] J. Lewis, "Usability testing," in *Handbook of Human Factors and Ergonomics*, G. Salvendy, Ed., chapter 46, pp. 1267–1312, John Wiley & Sons, New York, NY, USA, 4th edition, 2012.

[17] D. Manes, P. Green, and D. Hunter, "Prediction of destination entry and retrieval times using keystroke-level models," Technical Report UMTRI-96-37, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 1997, EECS-ITS LAB FT97-077.

[18] P. Green, "Estimating compliance with the 15-Second Rule for driver-interface usability and safety," in *Proceedings of the 43rd Annual Meeting on Human Factors and Ergonomics Society*, Human Factors and Ergonomics Society, CD-ROM, Santa Monica, Calif, USA, 1999.

[19] C. Nowakowski and P. Green, "Prediction of menus selection times parked and while driving using the SAE J2365 method," Technical Report 2000-49, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 2001.

[20] Calculation of the Time to Complete in-Vehicle Navigation and Route Guidance Tasks, SAE Recommended Practice J2365, 2002.

[21] S. Schneegass, B. Pfleging, D. Kern, and A. Schmidt, "Support for modeling interaction with automotive user interfaces," in *Proceedings of the the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '11)*, pp. 71–78, 2011.

[22] B. T.-W. Lin, P. Green, T.-P. Kang, A. Mize, A. Best, and K. Su, "Touch screen menu selection time and SAE J2365 predictions of them," Technical Report UMTRI-2012-9, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 2012.

[23] J. Sauro, "Margins of error in usability tests," 2009, http://www.measuringusability.com/test-margin.php.

[24] J.-S. Park, P. Green, and M. Alter, "Usability assessment of the mobis gen 3 navigation radio," Technical Report UMTRI-2010-23, University of Michigan Transportation Research Institute, Ann Arbor, Mich, USA, 2010.

[25] J. R. Lewis, "Sample sizes for usability tests: mostly math, not magic," *Interactions*, vol. 13, no. 6, pp. 29–33, 2006.

[26] R. A. Virzi, "Refining the test phase of usability evaluation: how many subjects is enough?" *Human Factors*, vol. 34, no. 4, pp. 457–468, 1992.

[27] J. Sauro and J. R. Lewis, *Quantifying the User Experience*, Morgan Kaufmann, San Francisco, Calif, USA, 2012.

[28] Y. Kinoe, "The VPA method: a method for formal verbal protocol analysis," in *Designing and Using Human-Computer Interfaces and Knowledge Based Systems*, G. Salvendy and M. J. Smith, Eds., pp. 735–742, Elsevier Science, Amsterdam, The Netherlands.

[29] Road Vehicles—Symbols for Controls, Indicators, and Telltales, ISO Standard 2575, 2010.

[30] Apple Inc., "OS X Human Interface Guidelines", Apple, Cupertino, Calif, USA, 2012, http://developer.apple.com/library/mac/documentation/userexperience/conceptual/applehiguidelines/OSXHIGuidelines.pdf.

[31] C. Nowakowski, P. Green, and O. Tsimhoni, "Common automotive navigation system usability problems and a standard test protocol to identify them," in *Proceedings of the the ITS-America Annual Meeting (CD-ROM)*, Intelligent Transportation Society of America, Washington, DC, USA, 2003.

[32] Techsmith, "Morae (software)," Techsmith, Okemos, Mich, USA, 2012, http://www.techsmith.com/morae.html.

[33] J. Reason, *Human Error*, Cambridge University Press, Cambridge, UK, 1990.

[34] S. A. Shappelland and D. A. Wiegmann, "The human factors analysis and classification system-HFACS," Final Report DOT/FAA/AM-00/7, Office of Aviation Medicine, Federal Aviation Administration, US Department of Transportation, Washington, DC, USA, 2000.

[35] T. S. Andre, H. Rex Hartson, S. M. Belz, and F. A. McCreary, "User action framework: a reliable foundation for usability engineering support tools," *International Journal of Human Computer Studies*, vol. 54, no. 1, pp. 107–136, 2001.

[36] K. Hornbaek and E. Forkjaer, "Comparison of techniques for matching usability problem description," *Interacting With Computers*, vol. 20, no. 6, pp. 505–514, 2008.

[37] R. Khajouei, L. W. P. Peute, A. Hasman, and M. W. M. Jaspers, "Classification and prioritization of usability problems using an augmented classification scheme," *Journal of Biomedical Informatics*, vol. 44, no. 6, pp. 948–957, 2011.

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2012, Article ID 807805, 12 pages
doi:10.1155/2012/807805

Research Article

A Neurofuzzy Approach to Modeling Longitudinal Driving Behavior and Driving Task Complexity

R. G. Hoogendoorn, B. van Arem, and S. P. Hoogendoorn

Department of Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands

Correspondence should be addressed to R. G. Hoogendoorn, r.g.hoogendoorn@tudelft.nl

Received 14 September 2012; Revised 20 November 2012; Accepted 4 December 2012

Academic Editor: Motoyuki Akamatsu

Copyright © 2012 R. G. Hoogendoorn et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Technological innovations can be assumed to have made the driving task more complex. It is, however, not yet clear to what extent this complexity leads to changes in longitudinal driving behavior. Furthermore, it remains to be seen how these adaptation effects can best be modeled mathematically. In order to determine the effect of complexity on empirical longitudinal driving behavior we performed a driving simulator experiment with a repeated measures design. Through this experiment we established that complexity of the driving task leads to substantial changes in speed and spacing. In order to provide insight into how complexity is actually related to changes in longitudinal driving behavior we introduce a new theoretical framework based on the Task-Capability-Interface model. Finally in this paper we take some first steps towards modeling of adaptation effects in longitudinal driving behavior in relation to complexity of the driving task through the introduction of a new neurofuzzy car-following model and based on the proposed theoretical framework. In this paper we show that this model yields a relatively good prediction of longitudinal driving behavior in case of driving conditions with differing complexity. The paper finishes with a discussion section and recommendations for future research.

1. Introduction

Technological innovations have increased the amount of information provided by road side and in-vehicle information systems dramatically. Systems such as adaptive cruise control (ACC), navigation systems, smart phones, in-vehicle and road side traffic information systems, and automatic lane control (ALC) have shifted drivers from being a controller of the driving task towards being a manager of information while driving. This shift in the role of the driver can be assumed to have made the driving task more complex.

This shift in the complexity of the driving task can be assumed to have a substantial influence on longitudinal driving behavior (i.e., driving behavior in the same lane) with presumably an influence on traffic safety. In this sense, the most compelling evidence of the influence of driving task complexity on traffic safety may stem from crash statistics. Violanti and Marshall [1] compared 100 randomly selected US drivers involved in a crash over the last two years with another 100 drivers who were not involved in crashes. The results indicated a risk ratio of 5.6 : 1 for drivers who talk more than 50 minutes per month on mobile phones. Furthermore, in Redelmeier and Tibshirani [2] 70 drivers who were mobile phone users and were involved in crashes with substantial property damage were studied. Redelmeier and Tibshirani reported the risk of a crash occurring being 4.3 times higher when using a mobile phone.

In addition, Tijerina et al. [3] investigated operating four commercial navigation systems while driving on an oval test track with traffic. In their experiment 16 drivers were involved. From this study it followed that when operating these devices drivers are often not looking at the road, and accident risk increased substantially. An important question is, however, to what extent complexity of the driving task actually influences empirical longitudinal driving behavior. Therefore, the first research question of this paper is "to what extent does complexity of the driving task influence empirical longitudinal driving behavior, represented by changes in speed v and following distance s?"

In order to answer this first research question we performed a driving simulator experiment with a repeated measures design among 25 participants. In a control condition, normal driving conditions were simulated, while in the experimental condition complexity was added to the driving task through the introduction of narrow lanes with road side concrete barriers. Using the obtained driving simulator data we statistically analyzed speed v and spacing s, providing us with an indication of adaptation effects in longitudinal driving behavior following a change in the complexity of the driving conditions.

In the influence of complexity of the driving task on longitudinal driving behavior, human factors may be assumed to play a substantial role. Examples of likely candidates in this context are mental workload (e.g., [4]), situational awareness [5], and static driver characteristics (e.g., age and driving experience). Drivers may, for example, be distracted by the information which is provided to them or distracted due to the need to operate the system and consequently pay less attention to what the lead vehicles are doing with an influence on their driving behavior (see, e.g., [6]). It is, however, not yet clear how these human factors are actually related to the adaptation effects in longitudinal driving behavior in relation to changes in the complexity of the driving task due to the previously mentioned shift in the role of the driver. In this sense in this paper, we introduce a theoretical framework based on the Task-Capability-Interface model by Fuller [7]. In this theoretical framework adaptation effects in longitudinal driving behavior, consisting of compensation and performance effects, come forth from interactions between driver capability and task demands.

Furthermore, it is not yet clear how the possible adaptation effects in longitudinal driving behavior following a change in the complexity of the driving task should be modeled mathematically. Modeling of driving behavior (i.e., car-following) is of high importance, as these mathematical models form the core of microscopic simulation models. These models are used to ex ante determine the influence of for example information systems and new vehicular technology on traffic flow operations, safety, and emissions. However, current mathematical models of longitudinal driving behavior insufficiently incorporate human factors. As we argued in the aforementioned that human factors may be assumed to play a substantial role in the occurrence of changes in longitudinal driving behavior in relation to complexity of the driving task, we conjecture that current mathematical models of longitudinal driving behavior are fundamentally inadequate to model these effects. Therefore, the second research question is "in which way can the influence of complexity of the driving task on longitudinal driving behavior best be modeled mathematically?"

In order to answer the second research question, we also take a first step towards modeling of the influence of complexity of the driving task on longitudinal driving behavior. We aimed at developing a new model based on a neurofuzzy logic modeling approach including human factors with structure and parametric learning using data derived from the driving simulator experiment. In this sense we determined the optimal model complexity and trained this model with optimal complexity. Finally we compared the predictions of the model with actual data.

The objective of this paper is therefore to experimentally determine the influence of complexity of the driving task on empirical longitudinal driving behavior as well as to develop and test a new mathematical model aimed at modeling the influence of complexity of the driving task on longitudinal driving behavior.

In the next section we provide a brief state of the art. In this section we present an overview of the available research on the influence of the complexity of the driving task on empirical longitudinal driving behavior. This section is followed by the introduction of a theoretical framework relating complexity of the driving task to longitudinal driving behavior based on the Task-Capability-Interface model by Fuller [7], followed by a discussion of mathematical modeling of longitudinal driving behavior in relation to complexity of the driving task.

The state of the art is followed by a presentation of the research method. In this section we provide an introduction to the driving simulator used in this study, followed by a description of the driving environment developed for this experiment, the characteristics of the participants, and the data analysis method used to statistically determine the effect of the complexity of driving conditions on empirical longitudinal driving behavior. In the following section we present the results with regard to the influence of driving task complexity on adaptation effects in empirical longitudinal driving behavior.

In the following section we describe the proposed approach to modeling the influence of the complexity of the driving task on longitudinal driving behavior, that is, the neurofuzzy modeling approach. Here, we describe the results of the structure and parametric learning, the training of the model, and validation of the model through a comparison of the predicted behavior versus the actual longitudinal driving behavior in relation to complexity of the driving task. This paper concludes with a discussion section in which conclusions are drawn from the results and recommendations for future research are provided.

2. State of the Art

2.1. Adaptation Effects in Empirical Longitudinal Driving Behavior in relation to Complexity of the Driving Task. Adaptation effects in longitudinal driving behavior due to complexity may be assumed to be governed by compensation and performance effects. Compensation effects entail the assumption that drivers regulate their driving behavior in order to compensate for any reduction in attention to the driving task. Very little research has been performed on compensation effects in relation to complexity of the driving task, as research is mainly focused on performance effects. However, research strongly suggests that drivers do engage in a range of conscious adaptations in their driving behavior in order to maintain a certain level of risk [8]. Numerous studies have shown that, at the operational level [9], drivers aim at reducing their risk level during the use of in-vehicle technology or due to external circumstances (e.g., adverse weather conditions). Compensation effects in longitudinal driving behavior may consist of speed reductions [8, 10, 11] and changes in the distance to the lead vehicle [12, 13].

In Haigney et al. [8] it was shown that mean speed decreased while participants were conversing on a mobile telephone. More recent research carried out in a driving simulator by Rakauskas et al. [11] also found that drivers' mean speed decreased and their speed variability increased while carrying out a conversation on a mobile phone. The aforementioned speed reductions could be the result of a modification of performance goals of drivers and the acceptance of a suboptimal level of performance. Finally in Strayer et al. [13] it was shown that conversing on a hands-free mobile phone while driving led to an increase in the distance to the lead vehicle.

Besides compensation effects due to changes in the complexity of the driving task, adaptation effects in driving behavior are also governed by performance effects (i.e., the quality of conducting the driving task). For example, in Brookhuis et al. [4] it was shown that significant effects were found in telephone conversations while driving on rearview mirror checking, the adaptation of speed to the speed of the lead vehicle, and braking in reaction to decelerations of the lead vehicle. In Makishita and Matsunaga [14] an experiment was performed in which reactions of drivers in various age groups were examined in order to assess the influence of driving task complexity. In their research experiments were performed on a simulated street in order to identify drivers with large reaction times and drivers whose reaction times are strongly affected by driving task complexity. The results show that a secondary task (mental calculations) increased average reaction times for all age groups. This secondary task increased the differences between age groups and individuals and increased differences in the drivers' individual performance. Reaction times especially of elderly drivers were affected substantially.

From the aforementioned it can be concluded that complexity of the driving task (e.g., through the addition of a secondary task) has a substantial influence on driving behavior through compensation effects and performance effects. However, as was mentioned before, research especially on conscious compensation effects following a change in the complexity of driving conditions is scarce and focuses mainly on mobile telephone conversations. It is therefore crucial to gain more insight into the influence of complexity of driving conditions on empirical longitudinal driving behavior. Furthermore, from the available research it does not become clear how these compensation and performance effects in driving behavior are actually related to complexity of the driving task. To this end in the next section we introduce a new theoretical framework based on the Task-Capability-Interface model by Fuller [7].

2.2. Introducing a Theoretical Framework Behavioral Adaptation to Changes in Complexity of Driving Conditions. In the previous section we discussed the available research on changes in driving behavior following a change in the complexity of the driving task. It was concluded that it may be assumed that complexity of the driving task has a substantial influence on longitudinal driving behavior, although research is quite scarce. The aforementioned, however, does not yet inform us how these changes in driving behavior are actually related to complexity of the driving task. To this end in this section we introduce a new theoretical framework.

In the proposed theoretical framework based on the Task-Capability-Interface model by Fuller [7], external circumstances (such as road design, weather interactions with other vehicles, roadside traffic management measures, and in-car technology) determine the complexity of the driving task. Complexity may have an influence on driver capability, moderated by driver characteristics, as well as on task demands. Therefore, in the proposed theoretical framework driving task difficulty comes forth from a dynamic interface between the demands of the driving task and the capability of the driver (see Figure 1).

In this context, Fuller [7] mentions that driver capabilities are restricted by biological personal characteristics of the driver as well as by driving experience. However, these capabilities due to biological personal characteristics (e.g., age, gender, and ethnicity) and driving experience alone do not determine the total temporal capabilities of the driver, as more dynamic variables play a substantial role as well. An example of a dynamic driver characteristic is activation level. Activation level is defined as the individuals degree of energy mobilization [15]. Activation level has been shown to have a substantial influence on driving behavior. For example, in Matthews et al. [16] the hypothesis was tested whether activation level is associated with driving performance. Eighty young adult participants performed a simulated test drive concurrently with a reasoning task. The data indicated that performance was characterized by adaptive mobilization of effort in order to meet the changing task demands. Drivers with a high activation level adapted to high levels of demand fairly efficiently, but were at risk of performance reduction when the task required little effort.

Another important determinant influencing driver capability is distraction. It can be assumed that in case of distraction (e.g., due to mobile telephone conversations while driving) driver capability will be reduced [4]. It can be assumed that especially driver distraction plays a substantial role in adaptation effects in driving behavior in relation to complexity of the driving task due to the shift in the role of the driver.

As is the case with driver capabilities, driving task demands are also related to a multitude of elements [7]. However, important elements in task demands are the elements over which the driver of the vehicle has direct

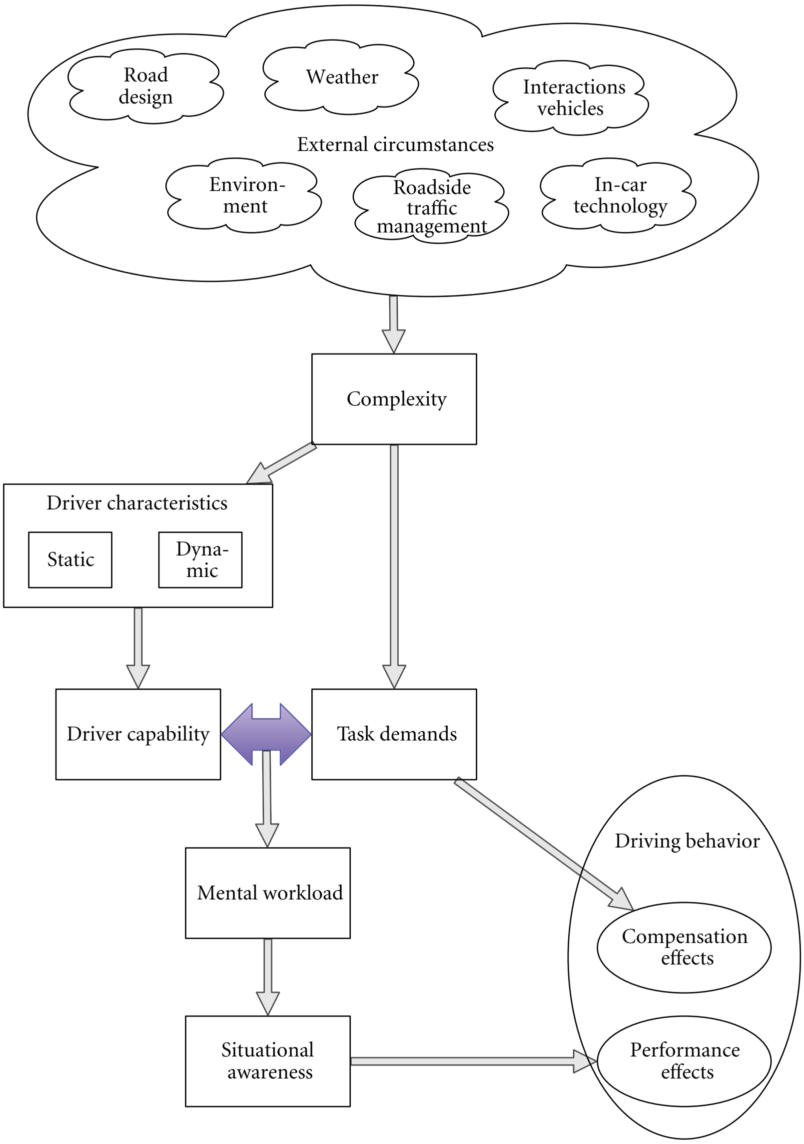


FIGURE 1: Theoretical framework of adaptation effects in longitudinal driving behavior in relation to complexity.

control. These conscious actions of the driver are in the ensuing referred to as compensation effects. Here, speed of the vehicle is clearly the most significant element: the faster a driver is moving, the less time is available to perceive stimuli, process information, and make decisions. As Taylor [17] regards the driving task as self-paced, driving task demand is in a fundamental way under the control of the driver through speed selection.

When drivers fail the driving task, a loss of control can be observed as a consequence. Thus, in essence, task difficulty is inversely proportional to the difference between task demand and the capability of the driver. According to Fuller [7], at the threshold where task demand begins to exceed the capability of the driver, a fragmented degradation of the driving task is to be expected. Fuller [7] continues by stating that with a static level of capability, any event that increases task demand will therefore reduce this critical difference, increase task difficulty, and potentially influence driving task performance.

In sum, changes in driver capability or task demands related to complexity (due to, e.g., in-vehicle systems or roadside traffic management measures) are expected to lead to compensatory changes in driving behavior (see also [4]). When these compensatory reductions are insufficient in order to balance task demand with driver capability, driving performance will suffer (performance effects).

2.3. Mathematical Models of Longitudinal Driving Behavior in relation to Complexity of the Driving Task. In the previous sections we discussed the available knowledge on the influence of complexity of the driving task on longitudinal driving behavior and proposed a new theoretical framework relating changes in driving behavior to complexity. This does not, however, yet inform us how adaptation effects in longitudinal driving behavior due to changes in the complexity of the driving task can best be modeled. To this end in the present section a brief overview is provided on current mathematical models of longitudinal driving behavior in relation to complexity of the driving task. In this context we will start with discussing a few often used models in which we will show that current models insufficiently incorporate human factors.

The GHR model [18] is perhaps the most well-known model of longitudinal driving behavior and dates from the late fifties and early sixties. The model is expressed in the following equation:

$$a(t) = cv^{m}(t)\frac{\Delta v(t-\tau)}{\Delta x^{l}(t-\tau)}. \tag{1}$$

In (1) a is the acceleration of a vehicle implemented at time t and is proportional to the speed v, relative speed Δv (speed difference with the lead vehicle), and relative distance to the lead vehicle Δx (distance headway) assessed at an earlier time $t - \tau$. In this equation τ represents the reaction time of the driver. Furthermore, in this equation m, l and c are the parameters to be determined. As acceleration a is dependent on relative speed Δv and relative distance Δx this model can be qualified as a stimulus-response model.

From this model it can clearly be observed that human factors are to a very limited degree incorporated. The only human factor that is incorporated in the GHR model [18] is a finite reaction time.

An alternative approach to car-following modeling was taken by Treiber et al. [19]. Their Intelligent Driver Model (IDM) was developed as the models developed up to this point had unrealistically small acceleration and deceleration times (e.g., in case of Bando et al. [20]) and because the more high fidelity models like the Wiedemann model [21] have too many parameters. Furthermore, Treiber et al. [19] conjectured that most models do not adequately incorporate traffic flow phenomena, such as traffic instabilities and hysteresis.

Acceleration in the IDM [19] is a continuous function incorporating different driving models for all speeds in freeway traffic as well as city traffic [22]. Besides the following distance Δx and speed v the IDM [19] also takes relative speed Δv into account. The IDM acceleration is given by

$$\begin{gathered} a = a_{\max}\left[1-\left(\frac{v}{v_0}\right)^{\delta}-\left(\frac{s^{*}(v,\Delta v)}{\Delta x}\right)^{2}\right], \\ s^{*}(v,\Delta v) = s_0 + vT + \frac{v\Delta v}{2\sqrt{a_{\max}\, b_{\max}}}. \end{gathered} \tag{2}$$

The expression combines a free flow acceleration regime $a[1-(v/v_0)^{\delta}]$ with a deceleration strategy $-a(s^{*}/\Delta x)^{2}$. The latter becomes relevant when the distance to the lead vehicle Δx is not significantly larger than the desired distance to the lead vehicle s^{*}. The free flow acceleration is characterized by free speed v_0, maximum acceleration $a_{\max}$, and the component δ. The component δ characterizes how acceleration decreases with speed.

The desired distance to the lead vehicle s^{*} is composed of a minimal stopping distance (jam distance) s_0, and a speed dependent distance vT. This corresponds to following the lead vehicle with a constant desired time headway T and a dynamic contribution which is only active in nonstationary traffic conditions. This implements an "intelligent" driving behavior that, in normal situations, limits braking decelerations to the maximum deceleration $b_{\max}$ [22].

A good step towards the incorporation of human factors in car-following models was taken by Tampere et al. [23] as in their macroscopic flow model finite reaction times, anticipation and driving style variations (i.e., attention level) were incorporated. In their gas-kinetic model the general law for the conservation of probability was transformed through use of the method of moments. Through this method the following macroscopic traffic flow model was derived:

$$\begin{gathered} \frac{\partial k}{\partial t} + \frac{\partial kV}{\partial x} = \left(\frac{dk}{dt}\right)_{\text{event}}, \\ \frac{\partial kV}{\partial t} + \frac{\partial(kV^{2}+k\theta)}{\partial x} = k\left\langle\frac{dv}{dt}\right\rangle_{v} + \left(\frac{dkV}{dt}\right)_{\text{event}}. \end{gathered} \tag{3}$$

In these equations k denotes the density, V denotes the speed, and θ denotes the distribution of speed. Driving style variations were implemented by characterizing an individual's state not only by the individual speed v and the distance to the lead vehicle, but also by the attention level a. Again the authors use the method of moments in order to obtain the speed dynamic equation [23]:

$$\begin{aligned} \frac{\partial A}{\partial t} + V\frac{\partial A}{\partial x} = {} & \left\langle\frac{da}{dt}\right\rangle_{v,a} \\ & + \frac{1}{k}\int_{a}\int_{v} a\left(\frac{d\rho}{dt}\right)_{\text{event}} dv \cdot da - \frac{A}{k}\left(\frac{dk}{dt}\right)_{\text{event}}. \end{aligned} \tag{4}$$

In (4) the first term on the right represents the effect of driver induced changes in the attention level A. The second term represents the effect due to events in the flow while the last term represents the redistribution of the total attention level A over the population k in case the density does not remain constant [23].

From this section it can be concluded that in its current form, current mathematical models such as the GHR model [18] and the IDM [19] insufficiently incorporate human factors (e.g., attention level, mental workload, and personal characteristics of drivers) and can therefore assumed to be less adequate in describing the effect of complexity of the driving task on longitudinal driving behavior and traffic flow operations. These elements are, to some extent, incorporated in the macroscopic traffic flow model by Tampere et al. [23]. As this is a macroscopic model, individual changes in driving behavior can be less adequately observed. Furthermore, this model lacks a thorough theoretical framework.

In Hoogendoorn et al. [24] we therefore proposed a new car-following model based on the theoretical framework discussed in this section. Besides the theoretical framework, this model was based on the previously discussed Intelligent Driver Model [19]. In the proposed model it was assumed that compensation effects and performance effects in longitudinal driving behavior following a change in the complexity of the driving task come forth from the difference between task driver capability and task demands at a certain time instant. The difference between task demands and driver capability $m_d(t)$ at time t is formulated as follows:

$$m_d(t) = m_t(t) - m_c(t). \tag{5}$$

In this equation $m_c(t)$ represents driver capability, while $m_t(t)$ represents task demands of a driver at time t. In the proposed model it was assumed that $-1 < m_d(t) < 1$. A driver will try to minimize the difference between driver capability and task demands, by exerting influence of those elements in driving behavior over which he has direct control (compensation effects). We assumed in the proposed model that a driver has direct control over maximum acceleration $a_{\max}$, maximum deceleration $b_{\max}$, free speed v_0, and desired time headway T (see also [25]). This is mathematically formulated as follows:

$$\begin{aligned}
a(t) &= \left(\left(-m_d(t)^3 a_{\max}\right) + a_{\max}\right) \\
&\quad \times \left[1 - \left(\frac{v(t)}{\left(-m_d(t)^3 v_0\right) + v_0}\right)^{\delta} - \left(\frac{s^*(v(t), \Delta v(t))}{\Delta x(t)}\right)^2\right], \\
s^*&(v(t), \Delta v(t)) \\
&= s_0 + v\left(\left(m_d(t)^3 T\right) + T\right) \\
&\quad + \frac{v(t)\Delta v(t)}{2\sqrt{\left(\left(-m_d(t)^3 a_{\max}\right) + a_{\max}\right)\left(\left(-m_d(t)^3 b_{\max}\right) + b_{\max}\right)}}.
\end{aligned} \tag{6}$$

In these equations it was shown that the contribution of maximum acceleration $a_{\max}$, maximum deceleration $b_{\max}$, free speed v_0, and desired time headway T to the IDM acceleration a is dependent on the difference between task demands and driver capability m_d. The effect of m_d on $a_{\max}$, $b_{\max}$, v_0, and T is assumed to be cubic.

Besides these conscious compensation effects, the adaptation of the Task-Capability-Interface model [7] also assumed that when a driver is unable to resolve the imbalance between driver capability and task demands, performance effects will occur. It is assumed that the difference between driver capability and task demands is related to performance effects with an inverted U-shaped function (see also [4]):

$$m_p(t) = -\left(\alpha m_d^2 + \beta m_d + \gamma\right). \tag{7}$$

In this equation m_p represents the performance effects in driving behavior due to complexity of the driving task at time t. Furthermore, in the equation α, β and γ are parameters.

Integrating performance effects into the Intelligent Driver Model [19] yields the following equations:

$$\begin{aligned}
a(t) &= \left(1 - m_p(t)\right)\left(\left(-m_d(t)^3 a_{\max}\right) + a_{\max}\right) \\
&\quad \times \left[1 - \left(\frac{v(t)}{\left(-m_d(t)^3 v_0\right) + v_0}\right)^{\delta} - \left(\frac{s^*(v(t), \Delta v(t))}{\Delta x(t)}\right)^2\right], \\
s^*&(v(t), \Delta v(t)) \\
&= s_0 + v(t)\left(\left(m_d(t)^3 T\right) + T\right) \\
&\quad + \frac{v(t)\Delta v(t)}{2\sqrt{\left(\left(-m_d(t)^3 a_{\max}\right) + a_{\max}\right)\left(\left(-m_d(t)^3 b_{\max}\right) + b_{\max}\right)}}.
\end{aligned} \tag{8}$$

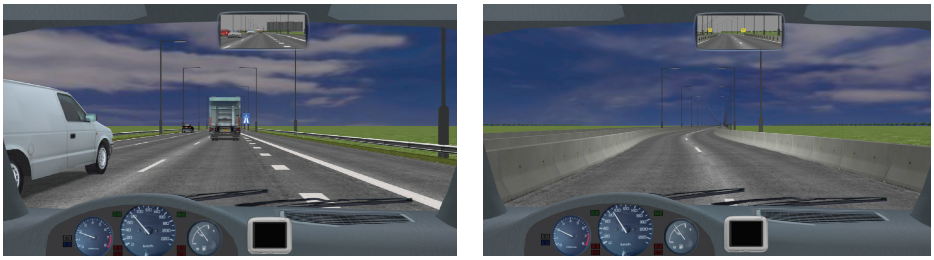
From Hoogendoorn et al. [24] it followed that the model showed quite well the influence of the difference between task demands and driver capability on longitudinal driving behavior and that the model was also able to provide a relatively adequate explanation for the so-called capacity funnel phenomenon as well as the influence of an optimal amount of information provision and information overload on driving behavior and macroscopic traffic flow operations. However, one of the main problems is to actually estimate changes in task demands and driver capability following a change in the complexity of the driving task. For instance, how does using a navigation system influence the balance between task demands and driver capability.

To determine these relationships in this paper we propose a new car-following model based on a neurofuzzy network approach. We chose this approach as a neurofuzzy network modeling approach allows for learning of the model structure and enables the establishment of relationships between stimuli and output variables. However, before presenting this model more insight is needed into the influence of complexity on empirical driving behavior. To this end in the following section the research method of the driving simulator experiment is presented.

3. Research Method

In the previous section we discussed the available research on changes in driving behavior following a change in the complexity of the driving task and proposed a new theoretical framework aimed at relating complexity to changes in driving behavior. The introduction of the theoretical framework was followed by a discussion of the available mathematical models of driving behavior in relation to complexity. We concluded that human factors are not incorporated in these models and can therefore be assumed to be fundamentally inadequate in describing adaptation effects following a change in the complexity of the driving task.

However, in order to be able to propose an adequate mathematical model of longitudinal driving behavior in relation to complexity of the driving task, more insight



(a) (b)

Figure 2: Driving environment developed for the purpose of the experiment. On the left the control condition is displayed, while on the right the experimental condition is displayed.

was needed into the extent to which complexity of the driving task actually influences empirical longitudinal driving behavior. To this end in the present section we present the research method. In this sense we present the experimental design, followed by an introduction into the used driving simulator and a presentation of the developed driving environment. This section is followed by a description of the research sample and the data analysis method.

3.1. Experimental Design. All participants participated in the experimental condition as well as in the control condition, rendering up a complete within-subjects design. Adaptation effects in longitudinal driving behavior, represented by adaptations in speed v and spacing s, were measured through registered behavior in the driving simulator at a sampling rate of 10 samples per second during both conditions.

3.2. The Driving Simulator and Driving Environment. The fixed-base driving simulator consists of three screens placed at an angle of 120 degrees, a driver's seat mockup, and hardware and software interfacing of this mockup to a central computer system. From the driver's seat the view consists of a projection of 210 degrees horizontally and 45 degrees vertically. The software was developed by ST Software.

For the purpose of the experiment, a driving environment was developed consisting of three segments. The first segment consisted of a short test drive through a suburban area to accustom participants to driving in a driving simulator and also to investigate whether the participants were prone to simulator sickness.

The other two segments were used in the experiment. These test trials took place on a virtual motorway with three lanes in the same direction. The length of the three segments combined was 9.45 km. In the control condition normal driving conditions with a medium density were simulated, while in the experimental condition narrow lanes and roadside concrete barriers were applied aimed at increasing the complexity of the driving task (Figure 2). Traffic with a medium intensity was simulated. The behavior of the other vehicles was derived from a pilot study and consisted of larger values of spacing s as well as a reduction in speed s.

3.3. Participants and Data Analysis Method. The research population consisted of 25 employees and students of Delft University of Technology (16 male and 9 female participants). The age of the participants varied from 22 to 54 years with a mean age of 29.68 years (SD = 6.93). Driving experience varied from 1 to 35 years with a mean of 9.6 years (SD = 7.50).

Adaptation effects were analyzed through a comparison of the indicators of longitudinal driving behavior (i.e., speed v and spacing s) between the control and the experimental condition using a paired samples t-test with a significance level of 0.05.

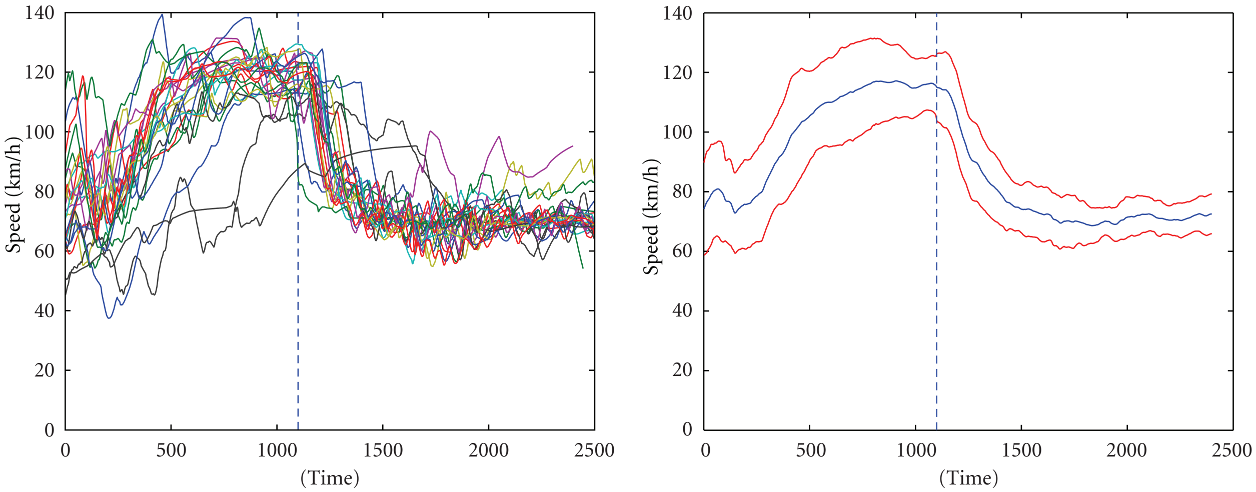
4. Results of Driving Task Complexity and Empirical Longitudinal Driving Behavior

In the previous section we presented the research method aimed at establishing the influence of complexity of the driving task (i.e., concrete barriers with narrow lanes) on empirical longitudinal driving behavior. In this context, the first research question was "to what extent does complexity of the driving task influence empirical longitudinal driving behavior, represented by changes in speed v and spacing s?"

From the paired samples t-test it followed that a significant reduction in mean speed in the experimental condition ($M = 77.80$, SD = 12.64) compared to the control condition ($M = 100.84$, SD = 16.12) could be observed, $t(24) = 4.23, P < .05$. As an illustration individual speeds and mean speeds along with the standard deviations are displayed in Figure 3.

Spacing s also showed a significant difference between the control and the experimental condition. In the control condition mean spacing amounted to 15.44 m (SD = 27.04), while in the experimental condition it amounted to 39.94 m (SD = 31.44). The difference between the two conditions was significant, as $t(24) = 9.88,\ P < .05$.

This leads to the conclusion that a change in the complexity of the driving task (i.e., narrow lanes and roadside concrete barriers) leads to substantial and significant compensation effects in longitudinal driving behavior, represented by a significant reduction in speed v and spacing s.



(a)

(b)

Figure 3: Individual speeds and mean speeds along with the standard deviation (expected value plus or minus the standard deviation) in case of normal and complex driving conditions. The dotted line indicates where the complex driving condition started.

5. Neurofuzzy Modeling of Longitudinal Driving Behavior in relation to Complexity of Driving Conditions

In the proposed theoretical framework we stated that complexity of the driving task has an influence on longitudinal driving behavior through task demands and driver capability. Indeed, it was established through the driving simulator experiment that complexity of the driving task has a substantial influence on empirical longitudinal driving behavior.

These adaptation effects in empirical longitudinal driving behavior due to a change in the complexity of the driving task reported in the previous section, however, do not inform us how these adaptation effects can best be modeled. In this sense the second research question was "in which way can the influence of complexity of the driving task on longitudinal driving behavior best be modeled mathematically?" In the state of the art we discussed several car-following models and concluded that these models, in general, can be assumed to be less adequate in describing longitudinal driving behavior in relation to complexity of driving conditions as they insufficiently incorporate human factors. In this sense we also briefly discussed the model introduced in Tampere et al. [23]. We argued that this model, as it is a macroscopic flow model, is less adequate in showing the influence of individual driving behavior and lacks a theoretical framework.

We also discussed the adaptation of the Intelligent Driver Model [19] as proposed in Hoogendoorn et al. [24]. Although this model shows relatively well the influence of the difference between task demands and driver capability as proposed in the theoretical framework, the main problem of this model is to actually estimate the influence of complexity of driving conditions on the difference between task demands and driver capability and ultimately longitudinal driving behavior. Therefore, in order to model adaptation effects in longitudinal driving behavior in relation to complexity of the driving task we propose a neurofuzzy modeling approach using a Takagi-Sugeno fuzzy architecture [26] and based on the proposed theoretical framework (difference in task demands and driver capability). In the next section we will start with an introduction into this approach.

5.1. *An Introduction into Neurofuzzy Logic Modeling.*

A Takagi-Sugeno fuzzy inference system [26] consists basically of a set of r rules, such as

$$\begin{aligned} &\text{if } x_1 \text{ is } A_1^1 \text{ and } x_2 \text{ is } A_2^1 \cdots \text{and } x_n \text{ is } a_n^1 \\ &\quad\text{then } y^1 = f^1(x_1, x_2, \ldots, x_n) \\ &\qquad\vdots \\ &\text{if } x_1 \text{ is } A_1^r \text{ and } x_2 \text{ is } A_2^r \cdots \text{and } x_n \text{ is } a_n^r \\ &\quad\text{then } y^r = f^r(x_1, x_2, \ldots, x_n). \end{aligned} \tag{9}$$

The antecedent of this equation is defined as a fuzzy AND proposition, where A_j^i is a fuzzy set on the jth premise variable defined by the membership function $\mu_j^i = \Re^n \to [0,1]$. The consequent is a function f^i with $i = 1, \ldots, r$ of the input vector $[x_1, x_2, \ldots, x_n]$. Through the fuzzy sets A_j^i the input is divided into smaller regions where the mapping is approximated by the models f^i. A weighted mean is used to recombine all the local representations in a global approximation:

$$y = \frac{\sum_{i=1}^{r} \mu^i y^1}{\sum \mu^i}. \tag{10}$$

In (10) μ^i represents the degree to which the ith rule is fulfilled. However, when the consequent is a linear model, for instance, in case of predicting accelerations, the system can be used to return a local linear approximation of a generic point of the input domain. Suppose that we have the input $\hat{x} = [\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_m]$. In this case (10) will return a linear approximation of $f_{\text{lin}}(\hat{x})$:

$$f_{\text{lin}}(\hat{x}) = \frac{\sum_{i-1}^{r} \mu^i \left(\sum_{j=1}^{n} p_{ij} \hat{x}_j + p_{i0} \right)}{\sum \mu^i}. \tag{11}$$

In the aforementioned traditional approach to fuzzy systems the membership functions and models are fixed according to prior knowledge (expert opinions). However, when this knowledge is (not yet) available, the components (given a certain data set) can be represented in a parametric form and the parameters are tuned through a learning procedure. In this case the fuzzy system turns into a neurofuzzy approximator [27]. In neurofuzzy systems, two types of learning are required, namely,

(i) structural learning;

(ii) parameter learning.

The first aims at finding a suitable number of rules and a proper partition of the input space (membership functions). Given an optimal structure, a neurofuzzy approximator searches for the optimal membership functions together with the optimal parameters of the consequent models.

5.2. Structure and Parameter Learning. We started with creating an input-output matrix with speed v, following distance s and relative speed Δv as inputs of the system derived from the driving simulator data discussed previously in this paper. Additionally we added a variable to the input-output matrix representing the difference between task demands and driver capability Co as input (as proposed in the previously discussed theoretical framework). In case of normal driving conditions a 0 was attributed to this variable while in case of a complex driving task (i.e., concrete barriers and narrow lanes) a 1 was attributed. This represents an increase in task demands following an increase in the complexity of the driving task (see also (5)). As output of the system we added acceleration a to the matrix. The input and output variables were all scaled between -1 and 1.

With regard to the structure of the neurofuzzy logic car-following models there may be a lot of different structure/parameter combinations which provide a feasible solution. We therefore aimed at finding the solution which provides the best performance in terms of generalization [28]. In the approach used in this paper we chose to use the number of rules as a measure of model complexity. To this end we adopted an incremental approach where different architectures with different levels of model complexity are assessed and cross-validated. The initialization of the architecture is provided by a hyperellipsoidal fuzzy clustering procedure [29]. In [29] it is proposed to cluster the data in the input-output domain through which a set of hyperellipsoids is obtained. This set can be regarded as a coarse representation of the input-output mapping.

Methods for initializing the parameters of a neurofuzzy system were derived from the procedure described in [30]. In the present paper we used the eigenvectors of the scatter matrix to initialize the parameters of the consequent functions f^i. Furthermore, we projected the cluster centers on the input domain to initialize the centers of the antecedents and adopted the scatter matrix in order to compute the width of the membership functions.

In the parametric estimation the best set of parameters was searched for by minimizing the sum-of-squares cost function J_M dependent solely on the training data set. As the model proposed in this paper is a linear model, the minimization procedure was decomposed into a least squares problem to estimate the linear parameters of the consequent models f^i [31] and a nonlinear minimization (Levemberg-Marquant) to find the parameters of the membership functions A_j^i [27] (see also (9)). In this context in this paper we used triangular-shaped membership functions of the antecedents. Mathematically, these membership functions can be formulated as follows:

$$\mu^i(x) = \prod_{j=1}^{n} \max\left(0, 1 - \frac{\left| x_j - c_j^i \right|}{b_j^i}\right). \tag{12}$$

The consequent model is mathematically formulated as follows:

$$y_i = \sum_{j=1}^{n} p_{ij} x_j + p_{i0}. \tag{13}$$

We determined the best model structure (related to the number of rules in the model) by gradually increasing the number of local models. Next we compared the different model structures in relation to their performance J_{CV} using a K-fold cross-validation [32]. We used a high proportion of the training data to determine the structure and the used method provided us with reliable estimates of the performance in generalization.

In Figure 4 the cross-validation versus complexity diagram is displayed. As was previously stated we chose the model with the lowest cross-validation error.

This leads to the conclusion that 6 rules provide the best model complexity:

$$\begin{gathered}
\text{if } x_1 \text{ is } A_1^1 \text{ and } x_2 \text{ is } A_2^1 \cdots \text{ and } x_n \text{ is } a_n^1 \\
\text{then } y^1 = f^1(x_1, x_2, \ldots, x_n) \\
\vdots \\
\text{if } x_6 \text{ is } A_1^6 \text{ and } x_2 \text{ is } A_2^6 \cdots \text{ and } x_n \text{ is } a_n^6 \\
\text{then } y^6 = f^6(x_1, x_2, \ldots, x_n).
\end{gathered} \tag{14}$$

As was mentioned before, with regard to the structure learning, we aimed at determining the membership functions, represented by the values of the centers c_j^i and the bases b_j^i (see (12)). In Tables 1 and 2 the values of the centers c_j^i and bases b_j^i are displayed, respectively. Note that i represents the number of the rule in the model, while j represents, respectively, the difference between task demands and driver

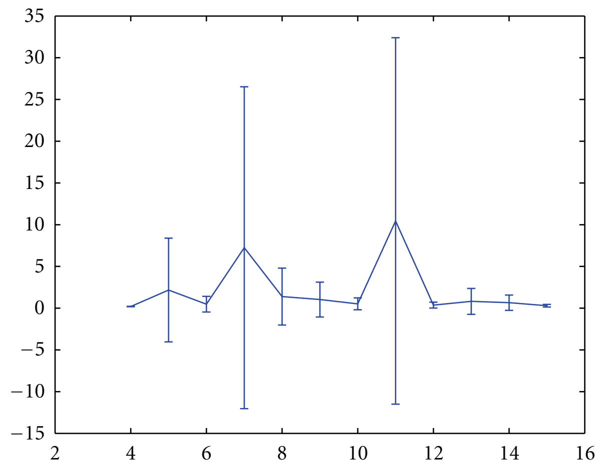


Figure 4: Cross-validation error versus complexity diagram. On the horizontal axis the number of rules is displayed, while on the vertical axis the cross-validation error is displayed.

Table 1: Results of estimation of c_j^i for the membership functions of the antecedents.

c_j^i	$j = 1$	$j = 2$	$j = 3$	$j = 4$
$i = 1$	0.2671	0.2970	0.1258	−0.0586
$i = 2$	0.0042	0.9617	0.2265	−0.2185
$i = 3$	0.0056	0.8374	0.2332	0.1199
$i = 4$	0.0065	0.4219	0.8657	0.8456
$i = 5$	0.0596	0.2392	0.0822	−0.1867
$i = 6$	0.0278	0.2524	0.0796	−0.2273

Table 2: Results of estimation of b_j^i for the membership functions of the antecedents.

b_j^i	$j = 1$	$j = 2$	$j = 3$	$j = 4$
$i = 1$	1.4656	1.3729	1.7483	1.9349
$i = 2$	1.9915	1.8448	1.5469	2.2208
$i = 3$	1.9887	1.5963	1.5334	2.2399
$i = 4$	1.9869	0.9362	1.7228	3.6481
$i = 5$	1.8806	1.4886	1.4330	2.1912
$i = 6$	1.9443	1.4620	1.8406	2.2723

Table 3: Results of estimation of the parameters p_{ij} and p_{i0} for the consequences.

p_{ij}	$j = 1$	$j = 2$	$j = 3$	$j = 4$	p_{i0}
$i = 1$	0.0000	5.1062	−22.2048	5.3034	4.9813
$i = 2$	0.0000	16.0963	−0.6176	2.5492	−11.7322
$i = 3$	0.8418	0.4348	−0.3756	0.3211	0.8418
$i = 4$	0.8095	2.7108	−0.3812	0.0757	0.8095
$i = 5$	0.0000	−2.9923	20.9578	−7.1156	2.4450
$i = 6$	0.0000	−1.5716	2.0211	−0.6674	1.2252

capability Co, speed v, distance to the lead vehicle s, and relative speed Δv.

Using the data from the driving simulator experiment we therefore aimed at estimating the parameters p_{ij} and p_{i0} for the consequences. In Table 3 the results of the estimation of these parameters are displayed.

5.3. Neurofuzzy Prediction of Acceleration.

The model structure and the value of the parameters, however, do not yet inform us to what extent this model actually provides a good prediction of accelerations in relation to complexity of the driving task. To this end we compared the accelerations of one driver given the established optimal complexity of the fuzzy architecture, the centers c_j^i and bases b_j^i of the membership functions, and the parameters p_{ij} and p_{i0} of the consequences to data from the driving simulator experiment.

In Figure 5 we provide two examples of the results for one individual driver. In the left graph the output of the fuzzy logic architecture is compared to the data from the driving simulator experiment under normal driving conditions while in the right graph the output is compared to the data from the driving simulator experiment under complex driving conditions. For convenience purposes in the graphs the scaled accelerations are shown. In the graphs the blue line represents the output accelerations from the fuzzy logic architecture, while the red line represents the scales' accelerations from the driving simulator experiment.

From Figure 5 it can be observed that overall the output of the fuzzy logic architecture resembles the actual accelerations of the driving simulator experiment fairly good, although in some cases the output of the architecture is somewhat less extreme. From an independent samples t-test it followed that the difference between the predictions and the driving simulator data was not significant ($P > .05$).

In this section we proposed a new car-following model able to model the influence of complexity of the task on longitudinal driving behavior using a neurofuzzy architecture and based on the Task-Capability-Interface model by Fuller [7]. We determined the optimal model complexity and determined the values of the centers and bases of the membership functions as well as estimated the parameters for the consequences. Finally we compared the output of the model to the actual data. We showed that the model provides a relatively good representation of accelerations under normal and in case of a complex driving task.

6. Discussion

Due to the technological innovations the amount of information directed at road users has increased substantially leading to shift in the role of the driver. Following this development the driving task is becoming more complex. Research has shown that complexity may be assumed to lead to adaptation effects in driving behavior although research is scarce and mainly focused on the use of mobile telephones while driving. Furthermore, a thorough theoretical framework on behavioral adaptation following a change in the complexity of driving conditions was lacking. Finally it was not yet clear how the adaptation effects in driving behavior in relation to complexity of the driving task can best be modeled mathematically.

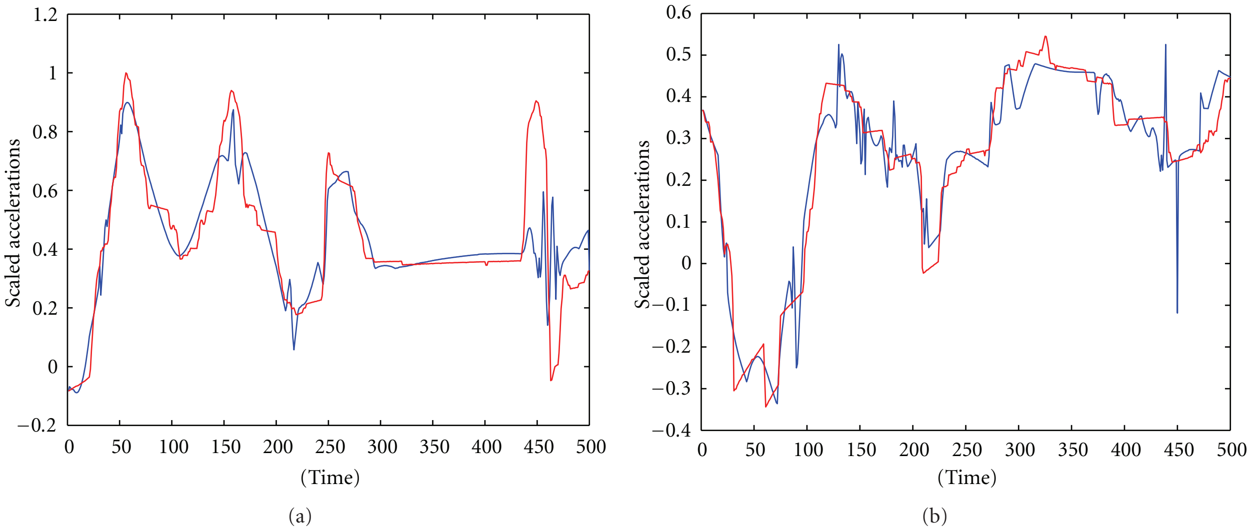


FIGURE 5: Scaled accelerations from the neurofuzzy acceleration model (blue line) and scaled accelerations from the driving simulator experiment (red line) for one driver. The left graph represents normal driving conditions, while the right graph represents complex driving conditions.

To this end we conducted a driving simulator experiment aimed at simulating the influence of a change in the complexity of the driving task on empirical longitudinal driving behavior (i.e., driving behavior in the same lane). Through this experiment we established that complexity of the driving task has a significant influence on speed v and spacing s. In our experiment speed v was significantly lower in the experimental conditions compared to the control condition, while spacing s was significantly higher in the experimental condition compared to the control condition.

In order to explain the adaptation effects in longitudinal driving behavior following a change in the complexity of the driving task we introduced a new theoretical framework based on the Task-Capability-Interface model by Fuller [7]. In this model adaptation effects in driving behavior come forth from the dynamic interface between task demands and driver capability following a change in the complexity of the driving task.

This framework was used as a basis for the proposed mathematical model of longitudinal driving behavior in relation to complexity. In the proposed mathematical model we used a neurofuzzy architecture with structure and parameter learning. We used this method as current mathematical models insufficiently incorporate human elements. The models which can be regarded as a good step towards the incorporation of human elements either lack a theoretical framework or are less adequate in determining the influence of the complexity of the driving task on task demands and driver capability (e.g., [24]).

In the context of the neurofuzzy approach we started with determining the optimal model complexity, after which we estimated the bases and centers of the membership functions of the antecedents and the parameters of the consequences. Finally we compared the proposed model to actual data derived from the driving simulator experiment. We showed that the accelerations under normal driving conditions as well as complex driving conditions resemble the actual accelerations fairly well.

From a human factors point of view the model is still quite limited as we only incorporated the difference between task demands and driver capabilities in the model. Also the difference between task demands and driver capabilities was fixed according to the conditions. We therefore recommend to extend the model through adding human factors, such as activation level, the level of distraction, age, and driving experience as input into the neurofuzzy architecture. This allows for a more adequate approximation of driver capability and task demands.

Furthermore, the model does not make an explicit distinction between compensation effects and performance effects following a change in the complexity of the driving task. Future research should therefore also focus on detailed analyses of driving behavior under different levels of driving task complexity. The results of these analyses can then be used as training data for the neurofuzzy logic car following model.

In this context, attention should be given to behavioral explanations of the rules incorporated in the model. This can be achieved through a more elaborate analysis of the model. Finally, a very limited dataset was used in order to determine the optimal model complexity and perform the estimations. In this sense we recommend to conduct future research in which a more elaborate data set is used, in which different kinds of complexity are incorporated.

Acknowledgment

The authors would like to thank the reviewers for their many useful comments and suggestions.

References

[1] J. M. Violanti and J. R. Marshall, "Cellular phones and traffic accidents: an epidemiological approach," *Accident Analysis and Prevention*, vol. 28, no. 2, pp. 265–270, 1996.

[2] D. A. Redelmeier and R. J. Tibshirani, "Association between cellular-telephone calls and motor vehicl collisions," *The New England Journal of Medicine*, vol. 336, no. 7, pp. 453–458, 1997.

[3] L. Tijerina, E. Parmer, and M. J. Goodman, "Driver workload assessment of rute guidance system destination entry while driving: a test track study," in *Proceedings of the 5th ITS World Congress (CD-ROM)*, Seoul, South Korea, 1998.

[4] K. A. Brookhuis, G. De Vries, and D. De Waard, "The effects of mobile telephoning on driving performance," *Accident Analysis and Prevention*, vol. 23, no. 4, pp. 309–316, 1991.

[5] M. R. Endsley, "Theoretical underpinnings of situation awareness: a critical review," in *Situation Awareness Analysis and Measurement*, 2000.

[6] R. Hoogendoorn, S. P. Hoogendoorn, K. Brookhuis, and W. Daamen, "Mental workload, longitudinal driving behavior, and adequacy of car-following models for incidents in other driving lane," *Transportation Research Record*, no. 2188, pp. 64–73, 2010.

[7] R. Fuller, "Towards a general theory of driver behaviour," *Accident Analysis and Prevention*, vol. 37, no. 3, pp. 461–472, 2005.

[8] D. E. Haigney, R. G. Taylor, and S. J. Westerman, "Concurrent mobile (cellular) phone use and driving performance: task demand characteristics and compensatory processes," *Transportation Research*, vol. 3, no. 3, pp. 113–121, 2000.

[9] J. A. Michon, "A critical view of driver behavior models: what do we know, what should we do," in *Human Behavior and Traffic Safety*, pp. 485–520, 1985.

[10] H. Alm and L. Nilsson, "The effects of a mobile telephone task on driver behaviour in a car following situation," *Accident Analysis and Prevention*, vol. 27, no. 5, pp. 707–715, 1995.

[11] M. E. Rakauskas, L. J. Gugerty, and N. J. Ward, "Effects of naturalistic cell phone conversations on driving performance," *Journal of Safety Research*, vol. 35, no. 4, pp. 453–464, 2004.

[12] A. H. Jamson, S. J. Westerman, G. R. J. Hockey, and O. M. J. Carsten, "Speech-based e-mail and driver behavior: effects of an in-vehicle message system interface," *Human Factors*, vol. 46, no. 4, pp. 625–639, 2004.

[13] D. L. Strayer, F. A. Drews, and W. A. Johnston, "Cell phone-induced failures of visual attention during simulated driving," *Journal of Experimental Psychology*, vol. 9, no. 1, pp. 23–32, 2003.

[14] H. Makishita and K. Matsunaga, "Differences of drivers' reaction times according to age and mental workload," *Accident Analysis and Prevention*, vol. 40, no. 2, pp. 567–575, 2008.

[15] W. B. Cannon, *Bodily Changes in Pain, Hunger, Fear, and Rage*, D. Appleton and Company, 1915.

[16] G. Matthews, L. Dorn, T. W. Hoyes, D. R. Davies, A. I. Glendon, and R. G. Taylor, "Driver stress and performance on a driving simulator," *Human Factors*, vol. 40, no. 1, pp. 136–149, 1998.

[17] R. G. Taylor, "Personality traits and discrepant achievement: a review," *Journal of Counseling Psychology*, vol. 11, no. 1, pp. 76–82, 1964.

[18] D. C. Gazis, R. Herman, and R. W. Rothery, "Analytical methods in transportation: mathematical car-following theory of traffic flow," *Journal of the Engineering Mechanics*, vol. 89, no. 6, pp. 29–46, 1963.

[19] M. Treiber, A. Hennecke, and D. Helbing, "Congested traffic states in empirical observations and microscopic simulations," *Physical Review E*, vol. 62, no. 2, pp. 1805–1824, 2000.

[20] M. Bando, K. Hasebe, K. Nakanishi, A. Nakayama, A. Shibata, and Y. Sugiyama, "Phenomenological study of dynamical model of traffic flow," *Journal De Physique I*, vol. 5, no. 11, pp. 1389–1399, 1995.

[21] W. Leutzbach and R. Wiedemann, "Development and applications of traffic simulation models at the karlsruhe institut fuer verkehrswesen," *Traffic Engineering and Control*, vol. 27, no. 5, pp. 270–275, 1986.

[22] A. Kesting, M. Treiber, and D. Helbing, "Enhanced intelligent driver model to access the impact of driving strategies on traffic capacity," *Philosophical Transactions of the Royal Society A*, vol. 368, no. 1928, pp. 4585–4605, 2010.

[23] C. Tampere, Arem, B. van, and S. P. Hoogendoorn, "Gas-kinetic traffic flow modeling including continuous driver behavior models," in *Proceedings of the Annual Meeting of the Transportation Research Board*, Washington, DC, USA, 2003.

[24] R. G. Hoogendoorn, B. Van Arem, S. P. Hoogendoorn, and K. A. Brookhuis, "Applying the task-capability-interface model to the intelligent driver model in relation to complexity," in *Proceedings of the Annual Meeting of the Transportation Research Board*, Washington, DC, USA, 2013.

[25] R. G. Hoogendoorn, *Empirical Research and Modeling of Longitudinal Driving Behavior Under Adverse Conditions*, TRAIL, 2012.

[26] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 15, no. 1, pp. 116–132, 1985.

[27] H. Bersini and G. Bontempi, "Now comes the time to defuzzify neuro-fuzzy models," *Fuzzy Sets and Systems*, vol. 90, no. 2, pp. 161–169, 1997.

[28] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, New York, NY, USA, 1995.

[29] R. Babuska and H. B. Verbruggen, "Fuzzy set methods for local modeling and identification," in *Multiple Model Approaches to Modeling and Control*, R. Murray-Smith and T. A. Johansen, Eds., pp. 75–100, Taylor & Francis, 1997.

[30] R. Babuska, *Fuzzy modeling and identification [Ph.D. thesis]*, Technische Universiteit Delft, Delft, The Netherlands, 1996.

[31] J. S. R. Jang, C. T. Sun, and E. Mizutani, *Neuro-Fuzzy and Soft Computing*, Matlab Curriculum Series, Prentice Hall, 1997.

[32] M. Stone, "Cross validatory choice and assessment of statistical predictions," *Journal of the Royal Statistical Society B*, vol. 36, no. 1, pp. 111–147, 1974.

Hindawi Publishing Corporation
International Journal of Vehicular Technology
Volume 2012, Article ID 876892, 6 pages
doi:10.1155/2012/876892

Research Article

A Comparative Analysis of Subjective Quality of the Mobility between a New Portable Electric Transportation Mode and Walking

Hiroyuki Ohta,[1,2] Haruyuki Matsumoto,[1] Daisuke Fukuda,[1] and Satoshi Fujii[3]

[1] *Department of Civil and Environmental Engineering, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo 152-8552, Japan*
[2] *Ports and Harbours Bureau, Ministry of Land, Infrastructure, Transport and Tourism, 2-1-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8918, Japan*
[3] *Department of Urban Management, Kyoto University, Kyoto Daigaku Katsura 4, Nishikyō-ku, Kyoto 615-8540, Japan*

Correspondence should be addressed to Hiroyuki Ohta, otaota812@yahoo.co.jp

Received 27 July 2012; Accepted 29 November 2012

Academic Editor: Motoyuki Akamatsu

Copyright © 2012 Hiroyuki Ohta et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To analyze the psychological impacts of the introduction of new portable electric transportation modes, we implemented an experiment using a personal mobile vehicle (PMV). We investigated its effects on 2 types of the subjective quality of mobility (SQM): *instrumental* aspects including "easiness" and "speed": and *affective* aspects including "enjoyment," "seeing scenery," and "enjoying the atmosphere." The result indicated that PMV might contribute to the improvement of the instrumental aspects of SQM, but walking was regarded as more preferable in terms of the affective aspects. The results suggest that such a new transportation mode could contribute to the improvement of subjective quality of mobility, if and only if it can be introduced in an appropriate situation.

1. Introduction

In the recent years, various personal mobility devices (PMDs) such as powered wheel chairs, scooters, and Segways have been developed with needs for high personal mobility in the aged society and for mitigating the environmental problem and so forth, since PMDs might be alternative travel modes to automobile in a short-trip travel. in Japan, especially the Segway has been sold for companies since 2006; "i-real" and "PIXY," which are new PMDs on a conceptual phase, were exhibited at the Tokyo Motor Show 2007. These situations imply that new PMDs might be introduced into our society in the near future.

PMDs might contribute to the improvement of personal mobility, but they could also cause some negative problems such as accidental risk, vehicles' congestion, and health depression. To cope with these problems, Rodier et al. [1] tested the safety of low-speed transportation modes in the pedestrian environment, and Litman [2] suggested an appropriate way to manage some PMDs on nonmotorized facilities including walkways, sidewalks, and paths. Kin et al. [3, 4] proposed evaluation items of PMDs considering the human factors in terms of safety and risk and suggested the road assignment for various PMDs It is necessary to ensure the safety of PMDs, or to consider the road space assignment before these new vehicles will come into our society. However, it is still unclear how PMDs would influence the quality of people's mobility. One can assume that there would be various psychological influences in terms of, for instance, comforts or enjoyments.

2. Quality of Mobility

2.1. Intrinsic Value of Travel. Travel is not only a derived demand for the activities but also an activity for the travel itself. That means travel is not always regarded as a "cost" to be minimized, but travel has also positive aspects of travel

Table 1: Specification of PMV (in a development phase).

Weight	17 kg
Maximum speed	6 km/h
Diameter of tire	15 cmm

itself like strolling or driving pleasure. Previous studies showed an importance to distinguish between travel for its own sake and travel for utilitarian purposes and to the possibility that the factors which influence these two categories of travel may differ significantly [5–7]. Anable and Gatersleben [8] examined the relative importance of instrumental and affective factors for work and leisure journeys. From their results, people tend to attach more importance to the instrumental aspects for work journey, while attach almost equal importance to the instrumental and affective aspects for leisure journeys. In addition the enjoyable or favorable travel increased the demand of one's travel according to the studies by Choo et al. [9] and Ory and Mokhtarian [10]. As to short-trip travel that walking is one of the most important transportation modes, Handy [11] indicated that the purpose of many walking trips was the walk itself rather than whatever destination was reached. Cao et al. [12] analyzed the influences of built environment factors and residential self-selection on both types of pedestrian travel: travel for its own sake (strolling) and travel for utilitarian purposes (walking to the store). These studies give us evidences that there would be a significantly difference between travel for its own sake and travel for utilitarian purposes.

From these implications, if PMDs will be introduced and people will come to use it instead of the other modes including walking, both two aspects of travel would be changed. In other words, PMDs could change not only the travel for utilitarian purposes (e.g., improvement of speed), but also travel for its own sake (e.g., driving/walking pleasure). Therefore, it would be important to investigate the psychological impact of PMDs on users for the assessment of introduction of PMDs into an urban transportation system.

In this study, to analyze the psychological impacts of PMDs, we implemented an experiment using "a personal mobile vehicle" (PMV). "PMV" is a new portable electric transportation mode, which is a kind of PMDs, in a development stage with the following: 3 characteristics; (1) it can offer comfortable and short travel with low electric power which is environmental friendly; (2) it can be used safely in the pedestrian area; (3) people can take it along easily in different transportation modes such as bus, train, and car [13]. The specification of PMV used in this study is shown in Table 1, and its virtual image is shown in Figure 1.

2.2. Two Types of Subjective Quality of Mobility (SQM). In this study, we investigated 2 types of subjective quality of mobility (SQM): *instrumental* SQM including "easiness" and "speed" of mobility and *affective* SQM including "enjoyment," "seeing scenery," and "enjoying the atmosphere" of mobility. We assumed that *instrumental* SQM might be more important when people perceived a travel as a derived demand, for instance, an access to a destination. On the other hand *affective* SQM might be more important when people enjoy a travel for its own sake, for example, a stroll around.

Figure 1: Virtual image of PMV.

3. Method

We implemented an experiment at Tokyo University in late February 2007. 30 students (15 males and 15 females) and 32 workers (15 males and 17 females) were gathered from Tokyo Institute of Technology and Tokyo University. At first participants tried to drive PMV with an instruction by a trainer in 5 minutes, and then, they drove it freely in 10 minutes. After that, they moved to another room, and then, answered a questionnaire about the quality of mobility. Note that all participants had never experienced PMV.

3.1. Process in the Questionnaire. In the questionnaire, 4 types of different situations were described and questions regarding SQM in the situations were asked, respectively. The 4 types of different situations were as follows (1) access to the nearest restaurant from home, (2) stroll in the shopping area, (3) stroll in the city park, and (4) move in the airport.

As to "access to the nearest restaurant from home," we defined the nearest distance as from 500 m to 1 km and then asked about the participants, frequency to the nearest restaurant that they actually use. If they do not use any restaurant near their home, they were asked not to answer the questions regarding this situation. Note that the restaurant does not include "fast food shop" and "pub." As to "stroll in the shopping area" and "stroll in the city park," we provided sample pictures and then asked the participants to image similar place before questions. As to "move in the airport," we asked the frequency in use of most commonly used airport, and if someone never used an airport, we asked them not to answer the question about "move in the airport."

3.2. Measures. In the questionnaire participants were asked 3 questions regarding SQM, (1) the importance of different types of SQM in the situation, (2) different types of SQM for different modes in the situation, and (3) attitudes toward travel in the situations by different modes. In this study,

Table 2: Evaluation criteria of SQM.

	Importance of SQM	SQM for each transportation mode
Easiness	How easy you can travel	I can travel easy
Speed	How fast you can travel	I can travel fast
Enjoyment	How you can enjoy when traveling	I can enjoy when traveling
Scenery	How you can see various sceneries when traveling	I can see various sceneries when traveling
Atmosphere	How you can enjoy the atmosphere when traveling	I can enjoy the atmosphere when traveling

Table 3: Means and standard deviations (SD) of the importance of SQMs in each situation.

	Easiness		Speed		Enjoyment		Scenery		Atmosphere	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Access to the restaurant ($N = 58$)	5.81	(2.19)	4.36	(2.32)	2.60	(2.09)	2.16	(2.27)	1.88	(1.97)
Move in the airport ($N = 58$)	6.38	(0.85)	6.10	(1.36)	1.60	(1.70)	1.52	(1.81)	1.71	(1.67)
Stroll in the shopping area ($N = 61$)	4.07	(2.50)	0.89	(1.53)	4.83	(2.07)	5.07	(1.80)	5.20	(1.89)
Stroll in the city park ($N = 61$)	2.52	(2.36)	0.39	(0.95)	5.54	(1.44)	6.23	(1.01)	5.97	(1.14)

walking and PMV were selected as competitive transportation modes because the fundamental concept of PMV is that it would be primarily used in the pedestrian area. Only in case of the analysis about the access to the nearest restaurant, we added an automobile for a transportation mode because people can use it only in this case and access to a destination faster than any other private vehicles such as a bicycle and a motorcycle.

3.2.1. Importance of SQMs. We measured the importance of different types of SQM by using 5 evaluation criteria (*easiness*, *speed*, *enjoyment*, *scenery*, and *atmosphere*; see Table 2 for definitions) in each situation. Among these 5 criteria, we regarded *easiness* and *speed* as instrumental SQMs, and the others (*enjoyment*, *scenery*, and *atmosphere*) as affective SQMs. I asked the participants "How important are the following 5 criteria in this situation?" with the following instructions: "(1) select the criterion which you think the most important and rate it as on 7-point; (2) select the criteria which you do not think important at all and rate it as on 0-point; (3) make ratings the other criteria on numerical 6-point scale with verbal endpoints (1 = a little important, 3 = important, 6 = very important)."

3.2.2. SQM for Different Transportation Modes. After asking the importance of SQM, we asked 5 types of SQM (*easiness*, *speed*, *enjoyment*, *scenery*, and *atmosphere*) for each transportation mode in the situation. By using items shown in Table 2, I asked participants to make ratings on numerical 5-point scale with verbal endpoints (1 = do not agree at all, 3 = neutral, 5 = strongly agree) for each transportation mode.

3.2.3. Attitude toward Travel. We then measured two kinds of attitudes toward travel for each transportation mode in the situation. In case of "move in the airport" and "access to the nearest restaurant," we measured "negative attitude toward travel" by asking the participants to make ratings on a numerical 5-point scale with verbal endpoints (1 = do not agree at all, 3 = neutral, 5 = strongly agree), as follows. "If possible, I do not want to make this travel." On the other hand, in case of "stroll in the city park" and "stroll in the shopping area," these travels might be regarded as enjoyable travels for their own sake. Therefore, we used different measures for these situations, that is, for "intention to continue travel" by asking the participants to make ratings on a numerical 5-point scale with verbal endpoints (1 = do not agree at all, 3 = neutral, 5 = strongly agree), as follows. "I want to continue this trip". Note that both types of attitudes toward travel were measured for "move in the airport" because this travel might be regarded as an enjoyable one as well as a derived demand to reach the destination.

4. Results

4.1. Differences in the Importance of SQM. Table 3 shows the means of the importance of SQMs for the 5 criteria in each situation. Table 3 indicates that in case of "move in the airport" and "access to the nearest restaurant from home," means of "*easiness*" and "*speed*" were relatively high, while "*enjoyment*," "*scenery*," and "*atmosphere*" were relatively low. That is, *instrumental* aspects are more important than *affective* aspects for these travels. On the other hand, in case of "stroll in the city park" and "stroll in the shopping area," means of "*enjoyment*," "*scenery*," and "*atmosphere*" were relatively high, while means of "*easiness*" and "*speed*" were relatively low. as to "*speed*" especially the, means were less than 1.0; therefore, "*speed*" has very little importance for these travels. In addition, the *instrumental* aspects were the highest and the *affective* aspects were the lowest in case of "move in the airport", "*atmosphere*" has the highest importance in case of "stroll in the shopping area," and "*scenery*" has the highest importance in case of "stroll in the city park."

4.2. Effects of Transportation Modes on SQM. Table 4 shows repeated measures analysis of variance (ANOVA) results revealing that there were statistically significant differences between transportation modes for every SQM except for

Table 4: ANOVA results for SQMs among the transportation modes.

	Easiness	Speed	Enjoyment	Scenery	Atmosphere
Access to the restaurant	$F(2,114) = 4.61^{*}$	$F(2,114) = 40.17^{**}$	$F(2,114) = 21.59^{**}$	$F(2,114) = 31.99^{**}$	$F(2,114) = 2.28$
Move in the airport	$F(1,57) = 53.17^{**}$	$F(1,56) = 59.27^{**}$	$F(1,56) = 12.82^{**}$	$F(1,57) = 9.38^{**}$	$F(1,57) = 13.25^{**}$
Stroll in the shopping area	$F(1,60) = 10.92^{**}$	$F(1,60) = 42.87^{**}$	$F(1,60) = 6.22^{*}$	$F(1,60) = 131.8^{**}$	$F(1,60) = 86.40^{**}$
Stroll in the city park	$F(1,60) = 25.36^{**}$	$F(1,60) = 109.8^{**}$	$F(1,60) = 5.36^{*}$	$F(1,60) = 109.3^{**}$	$F(1,60) = 59.66^{**}$

$^{*}P < .05$, $^{**}P < .01$.

Table 5: Means and standard deviations (SDs) of SQMs for each transportation mode.

	Easiness		Speed		Enjoyment		Scenery		Atmosphere	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Access to the restaurant ($N = 58$)										
Walking	3.41	(1.35)	2.55	(1.27)	3.40	(1.11)	3.95	(1.13)	3.26	(1.28)
PMV	4.02	(1.08)	3.84	(1.11)	4.17	(0.96)	2.72	(1.07)	3.07	(1.11)
Automobile	4.02	(1.29)	4.36	(1.05)	2.98	(1.03)	2.64	(1.10)	2.81	(1.07)
Move in the airport ($N = 58$)†										
Walking	2.45	(1.43)	2.25	(1.35)	2.81	(1.11)	3.22	(1.31)	3.24	(1.16)
PMV	4.29	(1.14)	4.21	(1.16)	3.47	(1.02)	2.59	(1.04)	2.59	(1.03)
Stroll in the shopping area ($N = 61$)										
Walking	3.48	(1.16)	2.33	(0.98)	4.21	(0.88)	4.51	(0.67)	4.33	(0.65)
PMV	4.15	(1.03)	3.69	(1.19)	3.79	(1.03)	2.77	(1.06)	2.95	(1.01)
Stroll in the city park ($N = 61$)										
Walking	3.18	(1.10)	2.03	(0.86)	4.33	(0.63)	4.80	(0.40)	4.66	(0.57)
PMV	4.21	(0.91)	3.98	(1.13)	4.00	(1.00)	3.23	(1.17)	3.28	(1.28)

$^{\dagger}N = 57$ for "speed" and "enjoyment" in case of "move in the airport" due to missing value.

Table 6: Means and standard deviations (SDs) of the negative attitude toward travel.

	Mean	(SD)
Access to the restaurant ($N = 58$)		
Walking	2.33	1.22
PMV	1.93	1.11
Automobile	2.64	1.22
Move in the airport ($N = 58$)		
Walking	3.59	1.24
PMV	2.33	1.25

Table 7: Means and standard deviations (SDs) of the intention to continue the travel.

	Mean	(SD)
Stroll in the shopping area ($N = 61$)		
Walking	3.80	0.98
PMV	3.44	1.07
Stroll in the city park ($N = 61$)		
Walking	4.23	0.74
PMV	3.54	1.21
Move in the airport ($N = 58$)		
Walking	2.29	1.03
PMV	3.19	1.21

"*enjoying the atmosphere*" in case of "access to the nearest restaurant from home." Table 5 shows the means of SQMs for each transportation mode in each situation. As can be seen Table 5, the means of *instrumental* SQMs such as "*easiness*" and "*speed*" for PMV and automobile were higher than that for walking, whereas, the means of *affective* SQMs such as "*scenery*" and "*atmosphere*" for walking were higher than that for PMV and an automobile.

As to "*enjoyment*," when *instrumental* SQMs are important such as "move in the airport" and "access to the nearest restaurant from home," means for PMV were higher than that for walking. On the other hand, when *affective* SQMs are important such as "stroll in the city park" and "stroll in the shopping area," means for walking were higher than that for PMV. In addition the means for an automobile were the lowest in case of "access to the nearest restaurant from home."

4.3. Difference in the Attitude toward Travel. Table 6 shows means of "negative attitude toward travel" and Table 7 shows means of "intention to continue the travel." Repeated measures ANOVA results showed that there were statistically significant differences between transportation modes for the "negative attitude toward travel" in both cases and for the "intention to continue the travel" in case of "stroll in the city park" and "move in the airport" ($P < .01$), while marginal differences in case of "stroll in the shopping area" ($P < .10$). As can be seen from Table 6, means of the "negative attitude toward travel" for PMV were lower than other modes in case of "move in the airport" and "access to the nearest restaurant from home." In addition the mean for the automobile was the highest. Table 7 indicates that the means of "intention to continue the travel" for walking were higher than that for PMV in case of "stroll in the city park" and "stroll in the shopping area," but the result was opposite in case of "move in the airport." To sum up, the attitude toward travel for PMV was better than walking in case of "move in the airport" and "access to the nearest restaurant from home," whereas walking was better in case of "stroll in the city park" and "stroll in the shopping area."

5. Discussion

In this study we investigated 2 types of SQM through the experiment using PMV. The data showed that participants thought the *instrumental* SQMs were important in case of "move in the airport" and "access to the nearest restaurant from home," while *affective* SQMs were important in case of "stroll in the shopping area" and "stroll in the city park." Thus the importance of SQM was different between situations. "Move in the airport" and "access to the restaurant from home" might be a derived demand whereas "stroll in the park" and "stroll in the shopping area" might have an intrinsic positive utility. That is, it can be considered that the former is an *instrumental*-purpose travel, and the latter is an *affective*-purpose travel.

As to SQM for each transportation mode, PMV could gain the advantage in the *instrumental*-purpose travel, while walking might dominate in the *affective*-purpose travel. As well, regarding "attitude toward travel," "intention to continue mobility," and "*enjoyment*" (one of SQMs), evaluations of PMV were better than walking in the *instrumental*-purpose travel, while evaluations of walking were better than PMV in the *affective*-purpose travel. attitude toward travel for "move in the airport" was especially worse than "access to the nearest restaurant from home," and this travel had the lowest importance of *affective* SQMs. This means that this travel may be cognitive as a cost to be minimized (i.e., a derived demand), and people may want to reduce it as much as possible. Therefore "*enjoyment*" for PMV, whose evaluation was higher than "*speed*" and "*easiness*," indicated higher value than that for walking. In addition "*enjoyment*" for an automobile was the lowest, so that this transportation mode may be unsuitable for this kind of travel. In case of "stroll in the shopping area" and "stroll in the city park", "intention to continue the travel" were higher than 3.0 (mid-value), and people attached importance to *affective* SQMs whereas they do not attached importance to *instrumental* SQMs. Thus these travels have positive value themselves, so that they do not expect to shorten these travels. Also note that "atmosphere" was most important in the shopping area, whereas "scenery" was the most important in the city park.

These results imply that the shift from walking to PMV might lead to improvement of the *instrumental* SQMs such as "*easiness*" and "*speed*"; therefore "*enjoyment*" was enhanced and the negative attitude was restrained. From this standpoint, PMV might be introduced into important *instrumental*-purpose travel such as "move in the airport", where people consider the travel as just only *means*. On the other hand, the *affective* SQMs such as "*scenery*" and "*atmosphere*" for PMV were lower than that for walking. in *affective*-purpose travel, "*enjoyment*" and "intention to continue the travel" for PMV were also lower. Therefore introduction of PMV especially might not be suitable in this kind of situation. That is, walking is the best way for the *affective*-purpose travel such as "stroll in the shopping area" or "stroll in the city park," where people considered the travel as its own value. These results, of course, depend on the experimental conditions in this study. Therefore, it is unclear when and where the same result can be obtained. For example, since participants in this study have never experienced PMV, the results might be different if they were well trained due to learning effects. However, based on the results that participants feel easier to drive PMV than walking (in Table 5), similar results to this study might be obtained even if they are well trained. In any case, the results in this study are expected to imply a certain tendency that PMV is suitable for the *instrumental*-purpose while walking is suitable for the *affective* purpose. As we mentioned in the introduction part, since people perceived their travel as not only just means to reach the destination but also enjoyment itself, a policy maker should take notice the possibility that PMV has both positive and negative effects on the meaning of mobility's quality.

References

[1] C. J. Rodier, S. A. Shaheen, and S. Chung, "Unsafe at any speed?: what the literature says about low-speed modes," Research Report UCD-ITS-RR-03-10, Institute of Transportation Studies, University of California, Davis, Calif, USA, 2003.

[2] T. Litman, "Managing diverse modes and activities on non-motorized facilities: guidance for practitioners," *ITE Journal*, vol. 76, no. 6, pp. 20–27, 2006.

[3] T. Kin, K. Shirasaka, and T. Terashima, "A study of evaluation items of personal travel modes for compatibility analysis," *Infrastructure Planning Review*, vol. 21, no. 3, pp. 769–780, 2004 (Japanese).

[4] T. Kin, N. Suzuki, and T. Terashima, "Diversification of personal travel modes and the problems of compatibility," *Infrastructure Planning Review*, vol. 20, no. 3, pp. 691–701, 2003 (Japanese).

[5] P. L. Mokhtarian and I. Salomon, "How derived is the demand for travel? Some conceptual and measurement considerations," *Transportation Research Part A*, vol. 35, no. 8, pp. 695–719, 2001.

[6] P. L. Mokhtarian, I. Salomon, and L. S. Redmond, "Understanding the demand for travel: it's not purely 'derived'," *Innovation*, vol. 14, no. 4, pp. 355–380, 2001.

[7] D. T. Ory and P. L. Mokhtarian, "When is getting there half the fun? Modeling the liking for travel," *Transportation Research Part A*, vol. 39, no. 2-3, pp. 97–123, 2005.

[8] J. Anable and B. Gatersleben, "All work and no play? The role of instrumental and affective factors in work and leisure

journeys by different travel modes," *Transportation Research Part A*, vol. 39, no. 2-3, pp. 163–181, 2005.

[9] S. Choo, G. O. Collantes, and P. L. Mokhtarian, "Wanting to travel, more or less: exploring the determinants of the deficit and surfeit of personal travel," *Transportation*, vol. 32, no. 2, pp. 135–164, 2005.

[10] D. T. Ory and P. L. Mokhtarian, "Modeling the structural relationships among short-distance travel amounts, perceptions, affections, and desires," *Transportation Research Part A*, vol. 43, no. 1, pp. 26–43, 2009.

[11] S. L. Handy, "Urban form and pedestrian choices: study of Austin neighborhoods," *Transportation Research Record*, no. 1552, pp. 135–144, 1996.

[12] X. Cao, S. L. Handy, and P. L. Mokhtarian, "The influences of the built environment and residential self-selection on pedestrian behavior: evidence from Austin, TX," *Transportation*, vol. 33, no. 1, pp. 1–20, 2006.

[13] H. Matsumoto, D. Fukuda, and S. Fujii, "A fundamental study on the intention of use toward the new portable electric transport mode," in *Proceedings of the 36th Meeting on Infrastructure Planning and Management*, 2007, (CD-ROM).