Author(s): Kujala, Tuomo; Saariluoma, Pertti

Title: Measuring Distraction at the Levels of Tactical and Strategic Control: The Limits of Capacity-Based Measures for Revealing Unsafe Visual Sampling Models

Year: 2011

Version:

Please cite the original version:

Research Article

Measuring Distraction at the Levels of Tactical and Strategic Control: The Limits of Capacity-Based Measures for Revealing Unsafe Visual Sampling Models

Tuomo Kujala[1] and Pertti Saariluoma[2]

[1] *Agora Center, University of Jyväskylö, P.O. Box 35, 40014 Jyväskylö, Finland*
[2] *Department of Computer Science and Information Systems, University of Jyväskylö, P.O. Box 35, 40014 Jyväskylö, Finland*

Correspondence should be addressed to Tuomo Kujala, tuomo.kujala@jyu.fi

The control theory of driving suggests that driver distraction can be analyzed as a breakdown of control at three levels. Common approach for analyzing distraction experimentally is to utilize capacity-based measures to assess distraction at the level of operational control. Three driving simulation experiments with 61 participants were organized to evaluate which kind of measures could be used to analyze drivers' tactical visual sampling models and the related effects of distraction while searching textual information on in-car display. The effects of two different text types were evaluated. The utilized capacity-based measures seemed to be insufficient for revealing participants' tactical behaviors or effects of text type. The measures of workload or performance did not indicate reliably the differences found between participants' visual sampling strategies or which text type is better for enabling safer task timing behaviors. Visual sampling measures did indicate effects of text type on participants' tactical abilities. Differences in participants' visual sampling strategies leading to different levels of systematicity in visual behaviors can explain the variances in visual sampling efficiency. Displays encouraging unsystematic glance allocation behaviors were found potentially the most distracting in relation to safe visual sampling of in-vehicle displays.

1. Introduction

Use of in-vehicle information systems (IVISs) and mobile devices for various purposes on road is increasingly popular. The growing availability of versatile information systems while driving underscores the importance of finding ways to assess in a valid way these systems' distraction-related safety risks on multiple levels [1]. This research should also guide in finding ways to design safer user interfaces for these systems.

Major advancements in the investigation of driver distraction have taken place under the paradigms based on the limited resources or information processing capacity of the human being (e.g., [2–4]). Dual-task experiments on driving, which are based on the premise that secondary tasks divert capacity or resources from the primary task, make it possible to uncover the possible consequences of a capacity limitation, such as deteriorated object and event detection

(e.g., [5–9]). Major parts of driver distraction research have focused on revealing this type of dual-task interference at the level of operational control [1]. From control-theoretical point of view [10], however, there are additional levels of control in driving. The level of operational control is the lowest level, at which the drivers control investment of their information processing resources at a timescale of milliseconds to seconds [1]. Above this level, there are the levels of tactical control, for example, distribution of tasks over time at a timescale of seconds to minutes, and of strategic control, for example, strategic planning and task prioritization at a timescale of minutes to weeks [1]. On the tactical level, distraction means failure in task timing, and, on the strategic level, it can be understood as inappropriate priority calibration [1].

It is known in expertise research that all the errors of experts cannot be explained in terms of limited capacity and

excessive workload [11]. Some of them seem to be connected, instead, to the information contents of mental representations, that is, mental contents [11]. From the capacity point of view, it is irrelevant whether the information in the limited processing system of a human operator is inclusive, correct, relevant, or making sense. The only critical aspect of that information is its complexity relative to the operator's capacity to perceive, attend to, or remember it in normal situations or in situations where the operators's capacity has decreased, for example, in the cases of low vigilance or mental underload. However, people might also misrepresent the situation they find themselves in and, consequently, they may err. In this way the nature of their mental contents can be used to explain why things may have gone in a suboptimal manner. This is why it is logical to apply content-based approach also when investigating human-computer interaction problems.

From this point of view, the limitations of the experimental research focusing on the level of operational control and the limited resources or capacity of the human operator are obvious. There are a lot of examples of studies with time pressured and not self-paced secondary tasks that surely reveal decreased driving performance and, thus, capacity limitations of the participants, but the external validity of the conclusions can be questioned (e.g., [12, 13], see also [14]). These types of experimental settings do not allow or investigate the tactical or strategic abilities of the participants in more realistic task environments.

A common denominator behind the measures utilized in these experiments is the logic of explanation that proceeds as follows: drivers make errors, because the situational workload exceeds the limited capacity of the driver (see Figure 1). In the following, we will refer to these types of measures as the capacity-based measures. The capacity-based measures are intended to measure either workload (e.g., subjective workload ratings, visual load, and physiological measures) or its performance effects (e.g., driving performance, task times, and reaction times). Thus, the measures seem to be alone inadequate for analyzing the tactical and strategic behaviors of the participants. Neither can they provide information on the qualities of drivers' ways of interacting with IVIS in the space of possible strategies [15]. For these reasons, the measures seem to be insufficient for discovering distraction effects and providing guidance on how to enhance IVIS interactions at the higher levels of control.

In this paper, we will try to illustrate the limits of the capacity-based measures and explore what kind of measures could be utilized to assess driver distraction and thinking at the levels of tactical and strategic control. While focusing on visual IVIS interaction, variance in glance durations has been used previously by [16–18] as a measure for visual sampling efficiency, that is, the efficiency of distributing visual attention over time between two tasks. Victor et al. [2] suggested that variance in glance durations towards IVIS increases as a function of in-vehicle task difficulty. However, visual sampling efficiency is not a capacity-based concept, because it does not describe the level of workload or the related performance effects (see [10]). Instead, it can provide information on the systematicity of the visual behavior
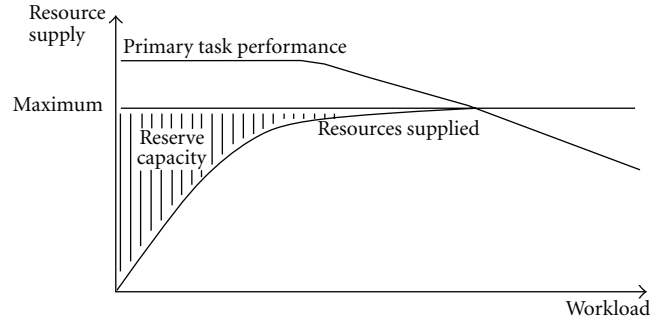

Figure 1: Capacity-based model of human dual-task performance (after [4]).

and thus, if the driver is able to utilize tactical or strategic thinking in the dual-task condition. Evidently, the variance in glance durations can increase by too short glances on display, inefficient for acquiring any useful information, or by overlong glances, which are potentially the most dangerous while driving [16, 20]. However, the variance in glance durations can also increase even if the driver is able to time glances to the IVIS efficiently in relation to the driving situation, but the driving demands are dynamic, as they often are in real traffic. This suggests that more exact measures that take the dynamics of the driving environment into account should be developed for visual sampling efficiency.

To evaluate if the measures of visual sampling efficiency could reliably reveal potential distraction-related risks of visual IVISs while allowing for tactical and strategic thinking, we organized three experiments with the help of an eye-tracking system and a series of self-paced visual secondary tasks on two different display designs in a dynamic driving simulation. We compared the explanatory power of the capacity-based measures to that of visual sampling efficiency measures regarding what types of measures can reveal how IVIS interaction designs should be enhanced for enabling safer task timing behaviors and strategies. The theoretical focus of investigation was on the limits of capacity-based measures for providing an exhaustive analysis of the significance of behavioral tactics and strategies in dual-tasking while driving.

2. Experiment 1

In the first experiment, we varied the properties of a text chapter displayed on an in-vehicle display in which the participants searched information while driving. We hypothesized that the capacity-based measures of subjectively assessed workload, task times, and driving performance lack the power of expression in determining which features of a visual display design are safer than others from the viewpoint of safe task timing behavior. In quantitative terms, there should not be significant effects of differences in visual displays on the capacity-based measures of subjective workload, total task times, or driving performance, although there would be effects of dual-task condition over single-task condition on this operational level of control. Instead, a detailed analysis of variances in visual sampling behavior

FIGURE 2: The driving simulation environment.

should reveal significant effects of display designs in the frequency of occasional overlong glances off road associated with safety risks [21] and failures in tactical task timing behavior. In addition, we hypothesized that this detailed analysis of variances in visual behavior with information-filled visual displays will reveal significant differences in individual visual sampling efficiencies that cannot be explained merely by capacity-based analysis.

2.1. Research Method

2.1.1. Participants. The 16 volunteers were recruited via public university e-mail lists. They included 9 women and 7 men between the ages of 20 and 33 (M = 24.3; SD = 3.9). They all had a valid driving license and lifetime driving experience from 2 to 100 thousand kilometers (M = 45.3; SD = 39.8). Eight of them were classified as experienced drivers ($\geq$30 000 km, 4 men, 4 women) and eight as novice drivers ($\leq$20 000 km, 3 men, 5 women). All the participants had a vision that was normal or corrected to normal. The experiment was conducted in Finnish with fluent Finnish speakers.

2.1.2. Materials. The tools used in the experiment included a driving simulator, consisting of a high-definition data projector, a computer, speakers, and a steering wheel with force feedback and pedals (see Figure 2). The use of the driving simulation in this case can be justified by the requirements of safety for the participants being subjected to demanding visual secondary tasks. We used the open-source Racer driving simulation software (http://www.racer.nl/). The software uses motion formulae from the actual engineering documents of the Society of Automobile Engineers. The car selected for the experiment was a Ford Focus RS with automatic gears, and it was adjusted for a realistic driving experience. The driving view included a speedometer and a tachometer just above the steering wheel.

For practice, the simulated road used was a track-like circuit, but the actual trials were driven on a two-lane rural road simulating a Polish country road with a speed limit of 50 km/h. Driving speed was kept within static limits between 40 and 60 km/h by instructions in our experiment. Also the width of the road, wind speed, and other possible factors affecting the position of the car were fixed. There was no other traffic on the road. The curvature of the road and visibility of the road ahead were the dynamic factors. Other equipments included a helmet-mounted SMI eye-tracking system with a 50 Hz sampling rate, video and audio capturing devices, questionnaires, and a computer for controlling the secondary task. A 17$''$ display was located 20 centimeters below the driving view and over 45° from the normal sight axis, on the right side of the participant.

2.1.3. Procedure. The experimental design consisted of a driving task with the driving simulator and a series of visual secondary tasks. The secondary tasks consisted of a spaced or compressed text, which made the discriminability of the text vary between groups (see Figure 3). The task design imitated situations in which the driver is reading a newsfeed or an e-mail message with an in-car Internet system while driving. The reading task was not the most typical of real-world in-car secondary tasks but the task could also be well related to in-vehicle visual search tasks that are common with, for example, modern point-of-interest or music track browsers. One of the examples of a news-related question in the display was, "How often disturbances caused by passengers happen in domestic flights?" The experiment included trials with and without the secondary task. The design was thus a within-subject design over dual-task condition and a between-subjects design over text type.

The instructions for the driving task were to keep the blue bonnet of the vehicle between the white lane markers and to keep the velocity of the vehicle between 40 and 60 km/h. In the secondary task, the participants were asked to answer orally questions located in the upper part of the display, drawn on text chapters below the question. The text and the related question changed after each correct answer. There was no time pressure in completing the secondary tasks, but the trial lasted as long as the participant took to complete five secondary tasks.

In the beginning of the experiment, general background information was collected from each participant. After this, the eye tracker was calibrated and overall instructions for the experiment were given. Prior to the trials, each participant was informed that the 10 most accurate participants in the driving tasks would be rewarded with movie tickets and that inaccuracy is defined by the total time spent outside the instructed areas (lane/speed zone). This was to make them prioritize the driving task. A practice driving task of about five minutes, consisting of driving around a track-like circuit, was performed without any secondary tasks. After this, the participants completed the trials with and without secondary tasks. The order of the trials was counterbalanced to eliminate learning effects from the driving performance data. The participants completed a reduced NASA Task Load Index (NASA-TLX) questionnaire (no weighting; [22]) after both trials. Before their dual-task trial, each participant was given, for practice, one secondary task without driving. Finally, the participants were shortly interviewed about their visual sampling strategies during the dual-task trial and about whether they were able to keep to their lane with peripheral vision (see [23]). The latter question is of great
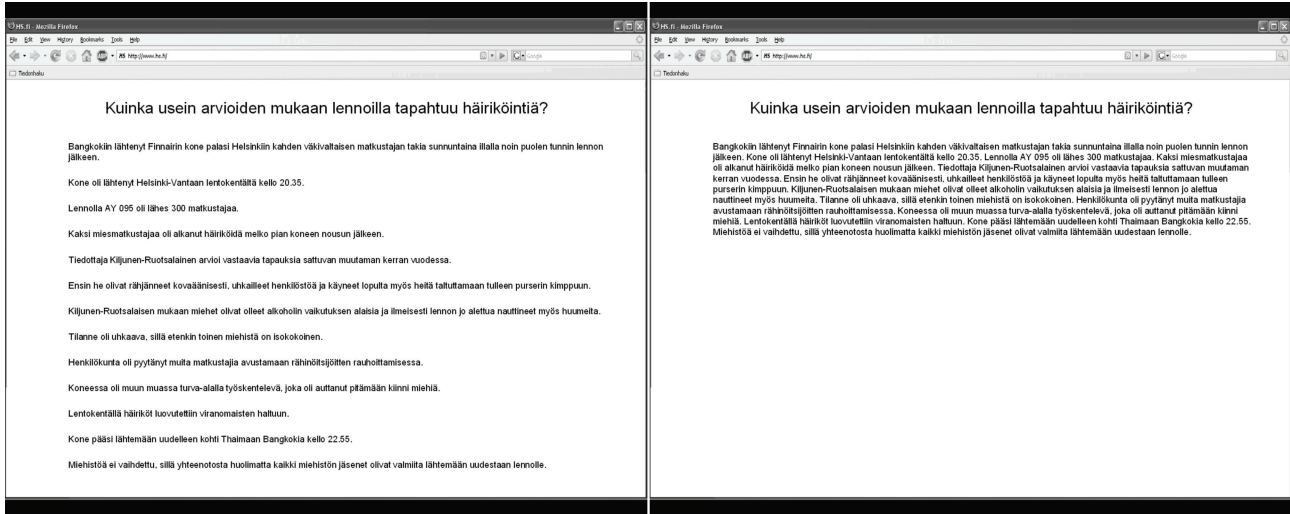
Figure 3: Two alternative ways to represent textual information on an in-vehicle display.

importance for the glance duration analysis and for the conclusions that follow [24].

2.1.4. Variables. The independent variables for analysis included the dual-task condition and the text type for the secondary task. The dependent variables included the total frequency and duration of lane excursions, which were defined to occur when the visible part of the car was over the lane markings for measuring deviations in lateral position; the total glance time; the total frequency of glances; the means, standard deviations, and maximum lengths of glance durations; the frequency of glances longer than 2 seconds, in total and while driving in curves; the NASA-TLX ratings. We also wanted to analyze whether task times [25] could reveal effects of text type. Glance durations of more than 2 seconds can be considered as unsafe in many circumstances [21, 26] (for similar measures, see [16]). Glances of over 2 seconds while driving in curves measured the participants' abilities to assess the difficulty of the driving situation and the extent to which they were able to adapt their task switching according to this information.

The controlled variables included driving experience and gender, which were balanced between the groups. Display properties other than the one varied were fixed between the different designs the same fonts (Arial, 12 pt), line spacing (single), and text locations. In addition, the point of the information searched for varied within text between tasks. The order of the trials was counterbalanced within groups.

2.1.5. Analyses. Mixed videos (25 frames per second) from the eye-tracking system and the driving scene were scored frame by frame for lane excursions, task times, and eye-movements with advanced video scoring software for behavioral research. A glance to the secondary task display was scored to begin at the frame the participant's gaze was off the road scene and to end at the frame with the gaze back in the road scene, following the SAE J2396 definition [27]. Other data included the interview notes of the experimenter.

Questionnaires were analyzed for means and variances within and between groups. Repeated measures ANOVA, two-tailed t-tests, and nonparametric Mann-Whitney U and Wilcoxon tests were used in order to find statistical significance and interaction effects in the results. Alpha level of 0.05 was used in statistical testing.

2.2. Results. The dual-task condition had a significant increasing effect on the mean frequency of lane excursions, from 3.38 to 12.63, $F(1, 14) = 19.38$, $P = 0.001$, and on the mean total duration of lane excursions, from 4.50 to 20.66 seconds, $F(1, 14) = 8.62$, $P = 0.011$. However, the text type did not have a significant effect on the frequency or duration of lane excursions. No interaction effects were found. Total task time did not indicate significant effects of text type, $t(14) = 0.380$, $P = 0.712$. Neither did the total task times correlate with the frequency or duration of lane excursions (Pearson's correlation).

The total duration or total frequency of glances at the secondary display during the tasks did not vary significantly between the text groups. Instead, the *Compressed Text* group had significantly greater means, $t(14) = 2.308$, $P = 0.039$, maximums $t(14) = 2.658$, $P = 0.030$, and standard deviations, $t(14) = 0.355$, $P = 0.010$ of glance durations at the text, compared to the *Spaced Text* group (see Table 1). They also made significantly more over-2-second glances to the display in total, $t(14) = 2.932$, $P = 0.015$, and while driving in curves, $t(14) = 2.966$, $P = 0.017$, than the *Spaced Text* group did. In the trial without secondary tasks there was a significant difference, $t(15) = 6.484$, $P < 0.001$, between the frequency of lane excursions committed in curves, $M = 3.19$, SD $= 1.87$, and the frequency of lane excursions committed on a straight road, $M = 0.06$, SD $= 0.25$, which indicates that driving in a curve was more demanding than driving on a straight road.

NASA-TLX measured the subjectively experienced demands of the tasks. There were no effects of text type in the results. Instead, excluding physical demand, there were

TABLE 1: Experiment 1: mean values of glance measures to the text display (standard errors).

	Spaced text ($n = 8$)	Compressed text ($n = 8$)
Total glance time (s)	209.92 (28.24)	249.82 (35.99)
Mean duration (s)*	0.97 (0.12)	1.46 (0.18)
Standard deviation of durations (s)**	0.59 (0.05)	1.33 (0.21)
Max duration (s)*	3.47 (0.44)	8.89 (1.99)
Total frequency of glances	230.50 (30.08)	174.50 (22.72)
Frequency of >2 s glances*	13.38 (3.88)	38.75 (7.74)
Frequency of >2 s glances in curves*	2.00 (0.68)	8.88 (2.22)

*Significant difference at 0.05 level; **significant difference at 0.01 level (t-test, two-tailed, equal variances not assumed).

differences between trials in all the reported scales: mental demand ($Z = -3.53$, $P < 0.001$), temporal demand ($Z = -3.41$, $P = 0.001$), effort ($Z = -3.37$, $P = 0.001$), performance ($Z = -3.37$, $P = 0.001$), and frustration ($Z = -3.04$, $P = 0.002$). Every participant in both groups rated the dual-task trial as more demanding on these scales than the trial without the secondary task.

In the short postexperiment interviews, it was found out that the participants had different visual sampling tactics. They tried to, for instance, allocate their visual attention to the display only while they were driving on a straight road or when the speed was easy to keep constant (no hills) or tried to maintain the lane position with their peripheral vision while reading the text. Eight participants reported that they tried to concentrate on reading the text and on maintaining the lane position at the same time with their peripheral vision, although they found it difficult, especially in curves.

2.3. Discussion. The dual-task condition had a significant effect on driving performance, which illustrates that capacity-based view of distraction as a breakdown of operational control can explain driving errors in our experiment. This was naturally expected, since extensive research on secondary tasks with driving has demonstrated this (e.g., [21, 26, 28]). However, text type did not have a significant effect on the frequency or duration of lane excursions, on total task times, or on the workload ratings. Instead, text type had an effect on the mean duration of glances, maximum glance durations, and variance of the glance durations. Finally, text type affected the frequency of overlong glances in total and in curves.

The first findings indicate that risky glance timing behavior is not in direct relation with capacity-based driving performance measures, total task times, or subjective workload ratings. As a matter of fact, the measures for driving performance, such as deviations in lane position, do not necessarily tell much about the differences of use risks between different secondary system designs, because the risks do not necessarily manifest themselves as driving errors or accidents

in experiments or in real traffic (e.g., [6, 24, 29]). Occurrence of a decrease in driving performance is a consequence of performance variability, the explanation of which requires a detailed description of the concurrencies of human and contextual variabilities and cannot be explained merely by the notion of an oversimplified conception of human performance [30, 31].

The other findings indicate that glance duration distributions are essentially affected by the differences in texts. The mean durations stayed on an acceptable level in both groups and are in line with Wierwille's et al. [32] visual sampling model. However, a more detailed analysis of glance durations can be indicative with respect to the efficiency of visual sampling behavior [16, 20]. It seemed that there were significant differences even between experienced drivers in their skills of visual sampling as measured by the variances in glance durations. In addition, the qualities of in-car visual displays, in this case, a difference as subtle as text spacing, were seen to have a clear effect on the visual sampling efficiency of the drivers. This means that detailed glance duration distributions are essential for analyzing risk factors in interaction between the driver and a visual IVIS while allowing for tactical and strategic thinking.

The data indicates that it is not sufficient to know that behaviors are different with respect to some gross capacity-based measure, such as the frequency of lane excursions between single-task and dual-task trials. This information does not yet tell what precisely should be changed in the interaction for enhancing safer dual-task behavior unless there is an obvious difference in the levels of workload or performance produced by two different display designs. Neither can capacity-based thinking explain the inter- and intraindividual differences in visual sampling behaviors, because participants' cognitive capacities are generally assumed to be basically equal between individuals and across situations (see [33]). Furthermore, the level of efficiency in visual behavior is a phenomenon that cannot be explicated by capacity-based explanations but rather requires analysis of the mental representations guiding visual attention. Accurate quantitative and qualitative information about how visual sampling behaviors and the respective individual mental representations are different would be required.

3. Experiment 2

The limitations of our first experiment included absence of other traffic on the road and, thus, low expectancy for unexpected hazardous events. These missing factors may have influenced the participants' visual sampling behavior [34]. However, the task of lane keeping by peripheral vision was evaluated as difficult by the participants, especially while driving in curves. This may have been due to the relatively large distance between the driving view and the secondary display [23]. Half of the participants were novice drivers, which could have magnified the effects of the text type (see [16]). The experimental design in the second experiment was improved to take these possible effects into account. Moreover, the analyses of glance durations in relation to the contents of the participants' search strategies were left open in

the first experiment. In the second experiment, we wanted to replicate and extend the findings of the first experiment, as well as to take the qualitative analysis of drivers' visual sampling and search strategies to a more detailed level. We wanted to find out if and how the measures of visual sampling efficiency relate to the contents of the found strategies.

Here, we hypothesized large variances in the visual sampling performance of the experienced participants and that the text type on a visual display would affect the participants' ability to utilize effective visual sampling and search strategies. Furthermore, we hypothesized that the quantitative capacity-based measures of performance would not provide us significant knowledge about the diverse safety-relevant effects of the differing visual display designs and, consequently, about how the interaction should be designed. Again, in quantitative terms, this means that (a) there should be significant differences between groups and large variances within groups with different visual displays in the measures of visual sampling efficiency, but (b) no significant effects of visual display design on capacity-based measures will be observable.

3.1. Research Method

3.1.1. Participants. The 18 participants (9 male and 9 female) were recruited from the public e-mail lists of the University of Jyvöskylö. This time, all participants were experienced drivers, in order to eliminate the effects of low driving experience in the results and to test more carefully the assumption that experienced drivers are generally capable of efficient visual time-sharing between driving and secondary tasks. The participants had a lifetime driving experience from 25 to 500 thousand kilometers, with a mean experience of 143.6 thousand kilometers (SD = 152.9), and they all drove a car on a weekly basis. The participants' ages were from 22 to 34 years, with a mean of 27.0 years (SD = 3.0). Again, all participants had a vision that was normal or corrected to normal, and the study was conducted in Finnish with fluent Finnish speakers.

3.1.2. Materials. The environment and the tools were the same as in the first experiment. A total of four secondary tasks similar to those in the first experiment were presented for the participants in the dual-task condition.

3.1.3. Procedure. The experimental design and procedure of the first experiment were replicated. However, improvements to the experimental design included oncoming traffic (four cars in preset points on the road) and an instruction to the participant to be aware of the possibility of unexpected events. This was to eliminate those overlong glances that would not be realistic in an actual driving environment and to encourage the participants to observe the driving environment for possible hazardous events as they would in real circumstances. The participants completed first the single driving trial and then the dual-task trial. The order of the trials was not counterbalanced this time, because the difference in driving performance measures between the trials was not our main interest and because we wanted to make the dual-task condition as similar as possible for every participant.

After the trials, an in-depth interview aimed at discovering the participants' visual sampling strategies was conducted by the experimenter. The interview included a question regarding whether the participants were able to remain in their lane with peripheral vision while focusing on the reading task.

3.1.4. Analyses. The measurements and analysis were nearly identical to those in the first experiment, with just some minor adjustments. The frequency and durations of lane excursions were analyzed for equal journey lengths between the two trials. The analysis of over-2-second glances in curves was done this time by an automatic script that compared the steering wheel movements recorded in the log file of the driving simulation to the synchronized data file scored from the eye-tracking video. The limit for driving on a curve was defined to be the absolute value of 0.60 or more of the steering wheel position in terms of the simulation's log file data, in which 0.00 was the calibrated center point. Variance in steering wheel position in the dual-task condition was analyzed as an additional capacity-based measure [35]. The interview recordings were analyzed after the experiments to seek commonalities and differences in the participants' descriptions of their visual sampling strategies, which were then used to classify the participants into the found types of strategies.

3.2. Results. Again, the dual-task condition had a significant increasing effect on the mean frequency of lane excursions committed, from 2.61 to 6.11, $F(1, 16) = 10.69$, $P = 0.005$. Significant differences were not found in the frequency or durations of lane excursions over the different text types. Neither did the text type have an effect on the variance in steering wheel position, $t(16) = 0.426$, $P = 0.676$. Total task time did not reveal significant effects of text type, $t(16) = 1.037$, $P = 0.317$, and did not correlate with the frequency or duration of lane excursions (Pearson's correlation).

The glance data analysis revealed that the *Compressed Text* group had significantly longer total glance times toward the secondary task display, $t(16) = 2.919$, $P = 0.017$. Additionally, the maximum durations of glances toward the display were longer, ($t(16) = 2.386$, $P = 0.041$), and the frequency of overlong glances in total, $t(16) = 2.300$, $P = 0.039$, as well as while driving in curves, ($t(16) = 2.245$, $P = 0.048$), was also significantly larger in the *Compressed Text* group (see Table 2). The distributions of glance durations in Figure 4 illustrate the effects of the text types. In total ($n = 18$), there were strong correlations between standard deviation of glance durations and frequency of over-2-second glances ($r = 0.904$, $P < 0.001$) and over-2-second glances in curves ($r = 0.941$, $P < 0.001$).

The participants rated the dual-task condition with NASA-TLX as mentally ($Z = -3.73$, $P < 0.001$), physically ($Z = -2.29$, $P = 0.022$), and temporally ($Z = -3.17$, $P = 0.002$) more demanding than the single-task condition. In their opinion, they also invested more effort ($Z = -3.30$, $P = 0.001$), felt their level of performance was lower ($Z = -3.73$, $P < 0.001$), and felt more frustrated ($Z = -3.38$, $P = 0.001$) in the dual-task condition. Again, there
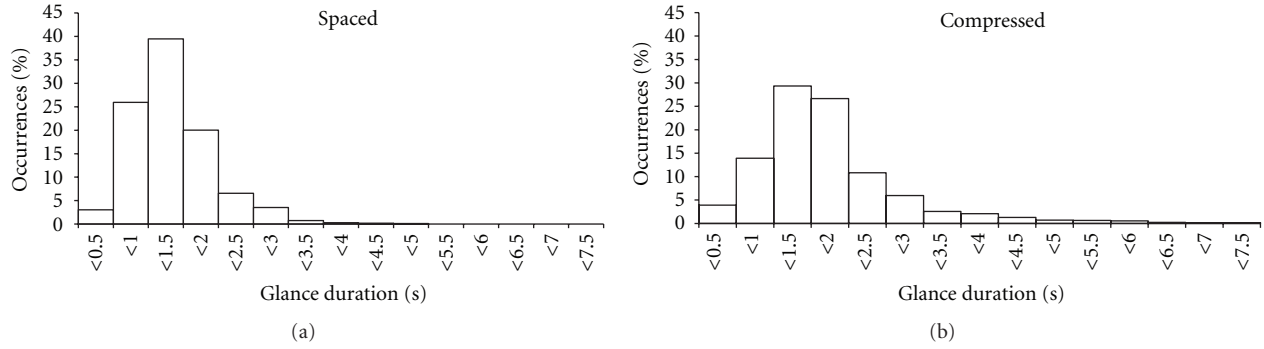
FIGURE 4: The distributions of glance durations by text type in Experiment 2.

TABLE 2: Experiment 2: mean values of glance measures to the text display (standard errors).

	Spaced text (n = 9)	Compressed text (n = 9)
Total glance time (s)*	157.08 (8.61)	255.08 (32.44)
Mean duration (s)	1.41 (0.12)	1.83 (0.21)
Standard deviation of durations (s)	0.49 (0.06)	0.97 (0.23)
Max duration (s)*	2.97 (0.34)	6.15 (1.29)
Total frequency of glances	116.44 (9.38)	139.67 (9.70)
Frequency of >2 s glances*	13.33 (5.15)	36.56 (8.68)
Frequency of >2 s glances in curves*	2.33 (1.64)	12.67 (4.30)

*Significant difference at 0.05 level (t-test, two-tailed, equal variances not assumed).

were no significant effects of text type in the subjective evaluations.

In the interviews, six main categories (S1–S6) were found for the visual sampling strategies developed by the participants during the dual-task trial (see Table 3 and Figure 5). When comparing the frequency of lane excursions and the used strategy, we found that the only two participants who drove perfectly (no lane excursions) also made no over-2-second glances toward the secondary task display. Both were in the *Spaced Text* group and followed the S1 strategy.

Six of the participants reported that they were not able or did not try to utilize peripheral vision for maintaining lane position while reading the text. The other 12 participants said they tried to use their peripheral vision to maintain their lane but it was difficult and possible only along the long straight parts of the road.

3.3. Discussion. There was a significant increasing effect of the dual-task condition on the frequency of lane excursions, but no significant effects of text type on lane excursions, steering wheel deviations, or subjective workload ratings. Neither did task time, suggested as a measure for distraction potential of IVIS [25], indicate any effects of text type.

Significant differences between the groups and large variances especially in the *Compressed Text* group in the visual sampling efficiency-related measures support the hypothesis of significant individual differences in visual sampling skills, even among experienced drivers.

The introduction of other traffic in Experiment 2 appeared to reduce, not increase, the interference with driving compared to Experiment 1. This suggests that, to some extent, drivers are able to reign in their driving performance when it becomes necessary to do so, presumably at the expense of the secondary task performance. In addition, the more experienced participants in Experiment 2 were likely more capable of perceiving the situational demands of driving correctly than the less experienced participants in Experiment 1.

In Experiment 2, difference between standard deviations in the groups' glance durations was not statistically significant ($t(16) = 2.021, P = 0.074$). However, there were strong correlations between this measure and the measures of overlong glances. These findings support the use of the measure for indicating differences between individuals' visual sampling skills with small samples [16–18], as well as a measure for drivers' abilities to combine a visual secondary task with driving in a safe and efficient way [36]. In addition, the measures of overlong glances in relation to the demands of the driving situation can reveal certain qualitative features of a display design that could induce unintentional visual distraction by the secondary task.

One could argue that the compressed text required more processing on the visual system than the spaced text, which would be a capacity-based explanation for the differences in the visual behaviors. However, this superficial explanation does not address our intent to understand the mechanisms on the tactical and strategic levels of control behind inter- and intraindividual variances in participants' visual behavior. The interviews revealed additional explanatory factors determining the efficiency of visual sampling than the mere discriminability of the text and that the information contents of the participants' mental representations of the dual-task situation can potentially explain the efficiency of visual sampling.

The results suggest that the more systematic the driver's visual sampling strategy, the more efficient the visual
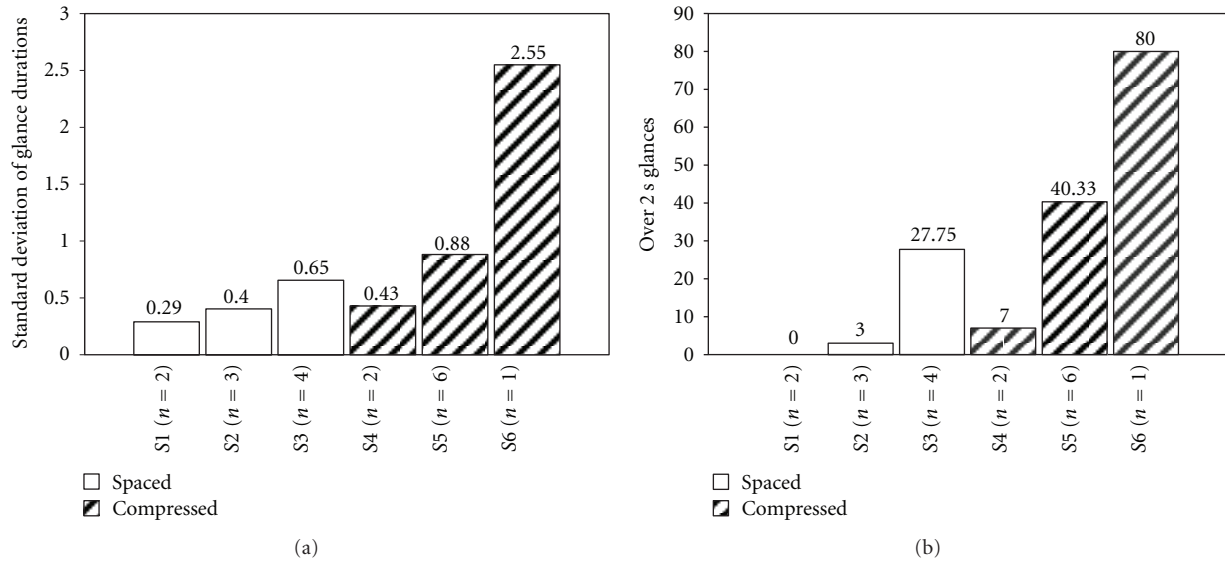
(a)

(b)

FIGURE 5: Visual sampling strategy versus mean values of visual sampling efficiency measures in Experiment 2.

TABLE 3: Subjectively Reported Visual Sampling Strategies of the Participants in the Experiment 2.

Text type	Strategy
Spaced	S1: The participant read from the beginning of each sentence, from top to bottom. Each sentence was read only so far as was necessary to comprehend whether the contents were relevant ($n = 2$).
	S2: The participant read the beginnings of the sentences, but jumped across sentences without a clear order ($n = 3$).
	S3: The participant read the beginnings of the sentences and jumped sentences, but also searched for relevant words from other parts of the sentences ($n = 4$).
Compressed	S4: The participant mentally cropped the text to sub areas by trying to find points of reference in the text, and searched for relevant words within the areas ($n = 2$).
	S5: The participant searched for relevant words everywhere without a clear order ($n = 6$).
	S6: The participant read every sentence until found the answer, from top to bottom ($n = 1$).

sampling performance (see Figure 5). It seems that the best visual sampling performance required a sufficiently inclusive and correct understanding of the dual-task situation's attentional demands and risks (see [1]). The compressed text did not seem to support the development of the presumably optimal strategy, S1. This strategy can be argued to be based on the information of systematic search behavior's importance in this context for keeping glance lengths to the display short and relatively static. We argue that, in the cases of the less systematic or otherwise less suitable visual sampling strategies and the resultant search behaviors, the gaze patterns and glance durations were controlled, at least partly, by the contingencies of the text, and not actively by the participant. This relates to the speed of finding where to start reading again with the next glance, how the time required to comprehend the meaning of a next sentence could be estimated, and how to adjust the corresponding glance duration and timing toward the visual display to the requirements of the driving situation. It seems that the key for a successful visual sampling was the capability of active control over one's glance allocation and glance lengths.

4. Experiment 3

In order to find additional quantitative support for our arguments derived from the results of the second experiment, we organized a third experiment in which we took a few of the visual sampling strategies found in the preceding interviews as our manipulated variables. We wanted to find out if the measures of visual sampling efficiency are really connected to the situation awareness and tactical decisions of the participants. We predicted that visual sampling (i.e., search) strategies can explain variances in visual sampling efficiency. In other words, participants who are trained to assume the most systematic visual sampling strategy, S1 (see Table 3), will perform significantly more efficiently in visual sampling than participants instructed in presumably less suitable strategies (S5: quick skim or S6: read all). These strategies were selected because they represent the three primary types of dual-task behavior, while the other strategies identified (S2–S4) are variants between these. We also hypothesized that the training will have a positive effect on driving performance.

4.1. Research Method

4.1.1. Participants. The 10 male and 17 female volunteer students (recruited through an e-mail service at the University of Jyväskylö) had a mean driving experience of 46.5 thousand kilometers (SD = 61.9), ranging from 2 000 to 250 000 kilometers. Their ages ranged from 19 to 32 years, with a mean of 23.1 years (SD = 3.2). All had a normal vision. As before, the experiment was conducted in Finnish with individuals fluent in Finnish. The participants were divided into three groups of 9, corresponding to three different visual sampling strategies (S1, S5, and S6). A between-subjects design was used to avoid any additional learning effects. Age, gender, and driving experience were balanced among the groups.

4.1.2. Materials. The tools and the driving simulation environment were the same as in Experiment 2. This time, however, all of the participants did secondary tasks with the spaced text type.

4.1.3. Procedure. First, participants practiced driving on the circuit-like track. After this they were trained for the visual sampling strategies with two secondary tasks. The participant was instructed for the strategy in the form seen in Table 3 (S1, S5, or S6). Between the two rehearsal tasks, the participant was inquired for success in the rehearsed strategy in the first task and the instructions for the selected strategy were repeated. Finally, participants drove the dual-task trial on the simulated rural road. A reaction task, consisting of braking to avoid colliding with a deer, was included for a further validation of visual behavior. While the reaction task was rehearsed once in the practice trial, the dual-task trial did not include the deer-related task, although the participants did not know this beforehand. All of the participants were rewarded with movie tickets.

4.1.4. Analyses. The measurements and analysis of visual behavior and driving performance were identical to those in the second experiment, with the exceptions of the single-dual-task comparison and the subjective workload measurements. The statistical multiple comparisons were made with one-way ANOVA with the Bonferroni correction and the alpha level .05. For controlling purposes, the participants were interviewed after the experiment to reveal if they reported failures in the rehearsed visual sampling.

4.2. Results. Glance data (see Figure 6) indicated that there were significantly longer total glance durations and a greater total frequency of glances in Group S6 (read all) than in Group S1 (systematic) (mean difference of glance durations (S6–S1) = 100.03 (18.88), $P < 0.001$, 95%Cl = 51.43 to 148.63; mean difference of the number of glances (S6–S1) = 55.2 (13.54), $P = 0.001$, 95%Cl = 20.4 to 90.1). Also Group S5 (quick skim) had significantly lower total glance durations (mean difference (S6–S5) = 74.02 (18.88), $P = 0.002$, 95%Cl = 25.42 to 122.62) and a lower total frequency of glances (mean difference (S6–S5) = 46.6 (13.5), $P = 0.006$, 95%Cl = 11.7 to 81.4) than Group S6.

The frequency of over-2-second glances in total in Group S6 was significantly greater than in Group S1 (mean difference (S6–S1) = 18.2 (6.6), $P = 0.031$, 95%Cl = 1.4 to 35.1). Also maximum glance durations indicated significant difference between Groups S1 and S6, in favor of Group S1 (mean difference (S6–S1) = 1.19 (0.45), $P = 0.045$, 95%Cl = 0.022 to 2.35).

Driving performance measures (Figure 6) revealed that the strategy S1 had a significant decreasing effect on the frequency (mean difference (S6–S1) = 11.7 (4.3), $P = 0.034$, 95%Cl = 0.7 to 22.6) and total duration of lane excursions (mean difference (S6–S1) = 18.79 (7.03), $P = 0.040$, 95%Cl = 0.69 to 36.88) as compared to strategy S6. There were no statistically significant differences between Groups S5 and S6 in these measures.

4.3. Discussion. the results support our hypotheses: The search strategies can explain at least partly the differences in visual sampling efficiency between groups. Consequently, training for the assumingly most efficient visual sampling (S1) led to a better overall driving performance compared to Group S6.

The strategies S1 and S5 (quick skim) were significantly faster for finding the answers than S6 (read all). A quick skim of text for relevant words without a clear order might seem to be a safe and fast way for finding information while driving. In fact, variants of this strategy were the most popular in the second experiment (see Table 3). However, drivers should be aware of the risks of this unsystematic behavior. Our results indicate that this type of unsystematic visual sampling seems to be potentially more dangerous while driving than the more systematic strategy S1. A possible explanation for this is that unsystematic gaze allocation can lead to uncontrolled glance lengths in relation to the demands of the driving situation. This was possibly also the case with strategy S6, although it first seems to be a systematic search strategy. There was a cost associated with interrupting the reading in the middle of a sentence: in this case the participants did not necessarily know where to continue at the next glance. Dual-tasking while driving with these unsystematic strategies is in many sense like a leap into unknown water.

Strategy S1 allowed the situational mental representations of the participants to include more accurate information regarding where to allocate gaze with the next glance and how long it will take to complete the following check. More predictable glance lengths can help in deciding when it is safe to switch attention to the display while taking into consideration the varying demands of the driving task. These aspects of situation awareness as well as tactical and strategic thinking seemed to be closely related to the utilized measures of visual sampling efficiency.

The results indicate some specific suggestions for design. IVIS visual display designs should not encourage unorganized skimming or other unsystematic search strategies. Thus, browsing Web pages and e-mail messages, as such, is potentially hazardous while driving because their structures are diverse and unpredictable. To encourage the construction of a proper strategy, for example, the first few words revealing the contents of a sentence at a time could be the maximum

limit for textual information provided to the driver (see also [36]).

A preliminary analysis of the gaze paths on the display suggests that additional metrics could be developed for measuring systematicity in visual search behavior, such as average saccade lengths between fixations during a single glance on the display. Further research with larger samples could validate whether unsystematic visual sampling strategies are in fact among the most popular and natural strategies in the driver population. There were no perfect performances in group S1 observed as in Experiment 2, which might suggest that the practiced strategy is not the most natural for all drivers or that they did not learn to utilize it perfectly in the time given.

5. Conclusions

The results of the three experiments indicate significant differences in visual sampling and tactical skills of drivers, even among experienced drivers, especially when the visual secondary task is demanding. Previously these differences have been found between novice and experienced drivers [16], between healthy drivers and drivers with cerebral lesions [17], and between young and aged drivers [18]. The current finding indicates that mere driving experience does not mean the driver is capable of safe visual sampling between the driving task and any possible secondary task.

A driver's visual sampling capabilities seem to be greatly affected by the qualities of the secondary task display (Experiments 1 and 2) and the contents of the driver's visual sampling strategy (Experiment 3). These results specify the visual sampling model of Wierwille et al. [32] for the part of parameters that can cause variations in the predictions of the basic model, particularly by lengthening glance durations over the safe limits by attention capture induced by a secondary task. Another possible factor, for example, would be the level of interest the driver has towards the information displayed, leading to varying levels of engagement in the secondary task [34]. On average, the participants tried to keep the durations of the glances within safe limits, which are often considered to be below 2 seconds in traffic [21]. However, more interesting than drivers' general behavior are the possibilities of occasional unintentional distraction [20]. Even one overlong glance can be dangerous in a wrong situation.

The results suggest that visual sampling efficiency in self-paced dual-tasks could be utilized as an externally valid and sensitive target for comparing distraction potentials of different visual IVIS display designs at the levels of tactical and strategic control. Individual variances in glance durations [16–19] and the durations of glances in relation to the task difficulty bandwidth of the driving situation [24] could be further developed as measures for visual sampling efficiency while multitasking with visual tasks. These measures can also provide a degree for task predictability and interruptability [1, 37], and they should work with relatively small sample sizes which is important for ensuring cost efficiency and rapidity of industrial testing practices. With the measures of visual sampling efficiency, it was possible to see that the spaced text type better supported visual sampling between the driving and search tasks than the compressed text type. However, this type of information search activity cannot be regarded as safe for most of the participants, at least without proper training of tactical skills.

The self-paced secondary tasks used in our experiments can be argued to be a more externally valid way to evaluate the effects of the tasks than would be the case if the participants were required to perform the tasks as quickly as possible (e.g., [12, 13]). It should be noted that people often are capable of circumventing their capacity limits with visual sampling strategies if they are allowed such a possibility and are able to perceive situational demands correctly [32, 38–40].

In conclusion, detailed analysis of glance duration distributions in self-paced dual-task conditions seems to be an important criterion for assessing safety risks related to visual lapses of tactical and strategic control in the context of visual IVIS. Driving performance measures are an important addition in this analysis. They can indicate decreased level of performance at the level of operational control compared to the single-task condition, but they do not necessarily reveal risky visual behavior that can, for example, make drivers miss relevant cues within a developing driving situation. Neither are decreases in driving performance easily traceable to certain characteristics of a display design.

On a general level, our experiments illustrate that the utilized capacity-based measures and concepts have limits in their power of expression. The traditional capacity-based interaction analysis seems to be insufficient with respect to the quality of interaction. Considering our results, it is unclear how interaction should be redesigned on the grounds of quantitative capacity-based measures, such as deviations in lateral position, subjective workload ratings, task times, and total or mean glance durations (i.e., visual load), for enabling safer tactical behaviors. The same result will presumably apply also to such capacity-based measures as reaction times or heart rate variability. These measures do not specify the contents of the drivers' interaction models.

Cognitive capacity is generally seen as a rather stable parameter over individuals and situations [33]. Therefore, it is difficult to explain strategy-based inter- as well as intraindividual variations—in this case, safety-relevant variations in visual sampling efficiency—solely in terms of cognitive capacity (see also [41]). The deeper analysis of the information contents of participants' situational mental representations, that is, mental contents [11, 42], indicated why the unsystematic glance allocation models are potentially dangerous. Mental contents can essentially vary over individuals and conditions and relate closely to situation awareness as well as to tactical and strategic thinking. The relationship between the contents of the found visual sampling strategies and the observed data indicated that the standard deviations of glance durations and the frequencies of over-2-second glances were associated with the information contents of the participants' situational mental representations, that is, the level of information regarding where to allocate glances. The frequency of over-2-second glances in curves was associated with the levels of situation awareness, that is, whether
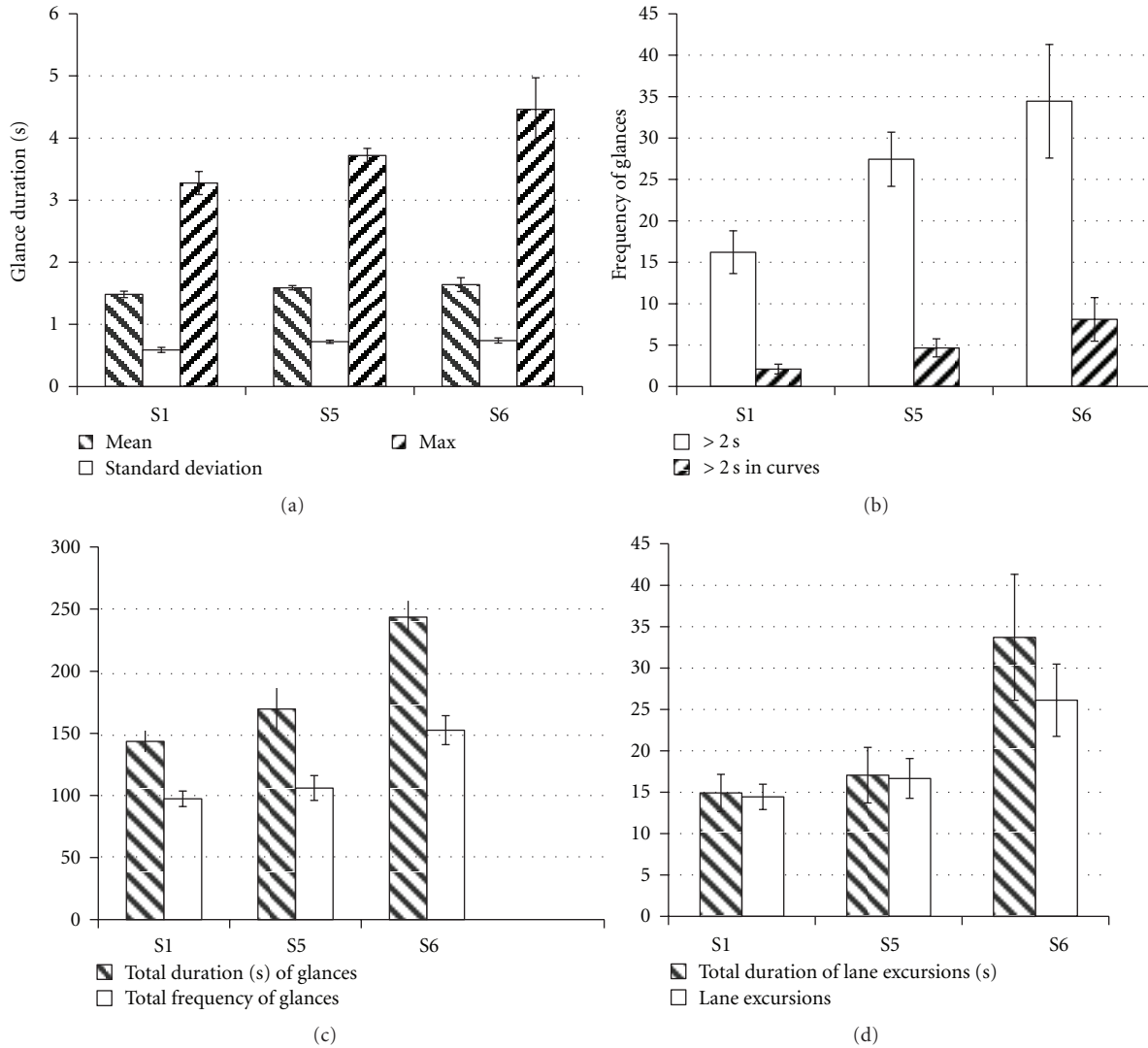
(a)

(b)

(c)

(d)

FIGURE 6: Experiment 3: mean values of glance measures to the text display and lane excursions by the strategies S1, S5 (quick skim), and S6 (read all). The bars represent standard errors.

the participants' situational mental representations included correct estimations on the required glance durations for keeping them in appropriate relation to the changing visual demands of the driving task.

More research is needed for revealing how much information is safe to display for the driver at one time and, more importantly, how this information should be presented to enhance the development of systematic and safe visual sampling behavior. Web pages or e-mail systems seem to be rarely optimized for searching textual information while driving. The investigation of the information contents of drivers' mental representations in self-paced dual-task situations can aid in revealing the exact mechanisms behind safety-relevant variances in visual sampling and task timings. This content-based approach needs not to be contradictory to the capacity-based. Instead, it is a complementary view; likewise the levels of control interact with each other [1]. This level of analysis would expand the focus of the research from workload to mental contents, which is, in the light of our

experiments, an important explanatory factor in distraction-related risk analysis at the levels of tactical and strategic control.

Acknowledgments

References

[1] J. D. Lee, M. A. Regan, and K. L. Young, "What drives distraction? Distraction as a breakdown of multilevel control," in *Driver Distraction: Theory, Effects, and Mitigation*, M. A. Regan, J. D. Lee, and K. L. Young, Eds., CRC Press, Boca Raton, Fla, USA, 2008.

[2] T. W. Victor, J. L. Harbluk, and J. A. Engström, "Sensitivity of eye-movement measures to in-vehicle task difficulty," *Transportation Research Part F*, vol. 8, no. 2, pp. 167–190, 2005.

[3] C. D. Wickens, "The structure of attentional resources," in *Attention and Performance VIII*, R. S. Nickerson, Ed., pp. 239–257, Erlbaum, Hillsdale, NJ, USA, 1980.

[4] C. D. Wickens, "Processing resources in attention," in *Varieties of Attention*, R. Parasuraman and D. R. Davies, Eds., pp. 93–102, Academic Press, Orlando, Fla, USA, 1984.

[5] P. Atchley and J. Dressel, "Conversation limits the functional field of view," *Human Factors*, vol. 46, no. 4, pp. 664–673, 2004.

[6] G. Jahn, A. Oehme, J. F. Krems, and C. Gelau, "Peripheral detection as a workload measure in driving: effects of traffic complexity and route guidance system use in a driving study," *Transportation Research Part F*, vol. 8, no. 3, pp. 255–275, 2005.

[7] Y. I. Noy, T. L. Lemoine, C. Klachan, and P. C. Burns, "Task interruptability and duration as measures of visual distraction," *Applied Ergonomics*, vol. 35, no. 3, pp. 207–213, 2004.

[8] D. D. Salvucci, "Predicting the effects of in-car interface use on driver performance: an integrated model approach," *International Journal of Human Computer Studies*, vol. 55, no. 1, pp. 85–107, 2001.

[9] W. W. Wierwille, "An initial model of visual sampling of in-car displays and controls," in *Vision in Vehicles IV*, A. G. Gale, I. D. Brown, C. M. Haslegrave, H. W. Kruysse, and S. P. Taylor, Eds., pp. 271–279, Elsevier, Amsterdam, The Netherlands, 1993.

[10] C. D. Schunn, M. C. Lovett, and L. M. Reder, "Awareness and working memory in strategy adaptivity," *Memory and Cognition*, vol. 29, no. 2, pp. 254–266, 2001.

[11] P. Saariluoma, *Foundational Analysis*, Routledge, London, UK, 1997.

[12] M. A. Just, T. A. Keller, and J. Cynkar, "A decrease in brain activation associated with driving when listening to someone speak," *Brain Research*, vol. 1205, no. C, pp. 70–80, 2008.

[13] T. B. Sheridan, "Driver distraction from a control theory perspective," *Human Factors*, vol. 46, no. 4, pp. 587–599, 2004.

[14] 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.

[15] D. P. Brumby, A. Howes, and D. D. Salvucci, "A cognitive constraint model of dual-task trade-offs in a highly dynamic driving task," in *Proceedings of the 25th SIGCHI Conference on Human Factors in Computing Systems*, pp. 233–242, ACM, New York, NY, USA, May 2007.

[16] A. S. Wikman, S. Haikonen, H. Summala, H. Kalska, M. Hietanen, and J. Vilkki, "Time-sharing strategies in driving after various cerebral lesions," *Brain Injury*, vol. 18, no. 5, pp. 419–432, 2004.

[17] L. Wierwille, S. Tijerina, T. Kiger, T. Rockwell, E. Lauber, and A. Bittner Jr., "Heavy vehicle driver workload assessment Task 4: Review of workload and related research," DOT HS Rep. 808 467, U. S. Department of Transportation, Washington, DC, USA, 1996.

[18] A. S. Wikman, T. Nieminen, and H. Summala, "Driving experience and time-sharing during in-car tasks on roads of different width," *Ergonomics*, vol. 41, no. 3, pp. 358–372, 1998.

[19] L. Tijerina, "Issues in the evaluation of driver distraction associated with invehicle information and telecommunications systems," Tech. Rep. S01-0926, Transportation Research Center, Ohio, USA, East Liberty, 2001.

[20] W. J. Horrey and C. D. Wickens, "In-vehicle glance duration: distributions, tails, and model of crash risk," *Transportation Research Record*, no. 2018, pp. 22–28, 2007.

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

[22] S. G. Hart and L. E. Staveland, "Development of NASA-TLX: results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds., pp. 139–183, Elsevier, Amsterdam, The Netherlands, 1988.

[23] J. Sodnik, C. Dicke, S. Tomažič, and M. Billinghurst, "A user study of auditory versus visual interfaces for use while driving," *International Journal of Human Computer Studies*, vol. 66, no. 5, pp. 318–332, 2008.

[24] W. J. Horrey, C. D. Wickens, and K. P. Consalus, "Modeling drivers' visual attention allocation while interacting with in-vehicle technologies," *Journal of Experimental Psychology*, vol. 12, no. 2, pp. 67–78, 2006.

[25] P. Green, "Estimating compliance with the 15-second rule for driver-interface usability and safety," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, pp. 987–991, Human Factors and Ergonomics Society, Baltimore, Md, USA, 2007.

[26] A. S. Wikman and H. Summala, "Aging and time-sharing in highway driving," *Optometry and Vision Science*, vol. 82, no. 8, pp. 716–723, 2005.

[27] P. Saariluoma, K. Kaario, K. Miettinen, and M. M. Mäkelä, "Mental contents in interacting with a multi-objective optimization program," *International Journal of Technology and Human Interaction*, vol. 4, pp. 43–67, 2008.

[28] SAE, *SAE J2396 Surface Vehicle Recommended Practice, Definitions and Experimental Measures Related to the Specification of Driver Visual Behavior Using Video Based Techniques*, Society of Automotive Engineers, Warrendale, Pa, USA, 2000.

[29] H. Summala, T. Nieminen, and M. Punto, "Maintaining lane position with peripheral vision during in-vehicle tasks," *Human Factors*, vol. 38, no. 3, pp. 442–451, 1996.

[30] S. W. Dekker, "Reconstructing human contributions to accidents: the new view on error and performance," *Journal of Safety Research*, vol. 33, no. 3, pp. 371–385, 2002.

[31] E. Hollnagel, "Human reliability assessment in context," *Nuclear Engineering and Technology*, vol. 37, pp. 159–166, 2005.

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

[33] N. Cowan, *Working Memory Capacity*, Psychology Press, New York, NY, USA, 2005.

[34] 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.

[35] J. D. Lee and D. L. Strayer, "Preface to the special section on driver distraction," *Human Factors*, vol. 46, no. 4, pp. 583–586, 2004.

[36] J. D. Hoffman, J. D. Lee, D. V. McGehee, M. Macias, and A. W. Gellatly, "Visual sampling of in-vehicle text messages: effects of number of lines, page presentation, and message control," *Transportation Research Record*, no. 1937, pp. 22–30, 2005.

[37] W. A. McDonald and E. R. Hoffman, "Review of relationships between steering wheel reversal rate and driving task demand," *Human Factors*, vol. 22, pp. 733–739, 1980.

[38] F. Cnossen, T. Rothengatter, and T. Meijman, "Strategic changes in task performance in simulated car driving as an adaptive response to task demands," *Transportation Research Part F*, vol. 3, no. 3, pp. 123–140, 2000.

[39] M. A. Recarte and L. M. Nunes, "Mental workload while driving: effects on visual search, discrimination, and decision making," *Journal of Experimental Psychology*, vol. 9, no. 2, pp. 119–137, 2003.

[40] C. D. Wickens and J. G. Hollands, *Engineering Psychology and Human Performance*, Prentice Hall, New Jersey, NJ, USA, 2000.

[41] R. E. Schlegel, "Driver mental workload," in *Automotive Ergonomics*, B. Peacock and W. Karwowski, Eds., pp. 359–382, Taylor ' Francis, London, UK, 1993.

[42] T. H. Rockwell, "Spare visual capacity in driving, revisited: new empirical results for an old idea," in *Vision in Vehicles II*, M . H. Freeman, C. M. Haslegrave, P. Smith, S. P. Taylor, and A. G. Gale, Eds., pp. 317–324, Elsevier Science, Amsterdam, The Netherlands, 1988.