Reducing drivers' mental workload by means of an adaptive man-machine interface

Walter Piechulla[a,*], Christoph Mayser[b], Helmar Gehrke[c], Winfried König[c]

[a]Institut für Experimentelle Psychologie, Universität Regensburg, now at Center for Traffic Sciences, Universität Würzburg

[b]BMW Group, Forschung und Technik

[c]Robert Bosch GmbH, FV/SLN Research and Development

ABSTRACT

Modern in-vehicle information and communication devices are changing the nature of the driving task. Drivers take it for granted that they are able to divide their attention between the primary task of driving and secondary tasks like monitoring information displays or using mobile phones. While it is commonly accepted that driver information overload can compromise traffic safety, attempts to introduce attention management within the vehicle are nowadays limited to restrictive decisions by legislative bodies. In an increasing number of countries, the use of hands-free phones is enforced by law. In some countries, the use of phones while driving is prohibited altogether. We argue that there is a more intelligent solution to the information overload issue, namely an adaptive man-machine interface that filters information presentation according to situational requirements. We implemented such a filter as a projective real-time computational workload estimator which is based on the assessment of traffic situations detected from an onboard geographical database. Workload estimates are refined by data from sensors that monitor the traffic environment and variables of driving dynamics. The prototype system is operational in a demonstrator vehicle. Whenever the workload estimate exceeds a threshold value, incoming telephone calls are automatically redirected to the telephone mailbox without notifying the driver. An evaluation field experiment that employed objective and subjective methods of assessing workload yielded promising results in terms of the possibilities of reducing workload by means of the adaptive interface. The results are in favour of the idea of a futuristic, situation-aware vehicle which has the potential to enhance comfort and safety while driving.

Keywords: Man-machine interface; Attention; Mental workload; Traffic; Phone

* Corresponding author. Tel.: +49-931-9709940

E-mail address: piechulla@psychologie.uni-wuerzburg.de

1. INTRODUCTION

Managing drivers' attentional resources is still a challenge for researchers and practitioners in the field of man-machine interaction. According to Verwey (1993), driver inattention plays a role in about 30% to 50% of all accidents. Sprenger (1999) even assumes that distraction from the main driving task is one of the most important causes of accidents.

At the same time, the number and variety of in-car information and communication systems affordable to a wider range of users is continuously increasing. The classic dashboard is being refitted to include navigation systems; in addition, drivers communicate by mobile phone and receive or even compose and send texts using the short message system (SMS). More recently, Internet portals have been introduced which offer information specially "tailored" to the needs of the driver. These services typically include email, parking information, and directory assistance, all of which are intended for use while driving.

On one hand, these new devices and services have the potential to enhance comfort, efficiency and even safety of the driving task, if they are used sensibly. On the other hand, however, ergonomists and psychologists are warning against distraction effects and information overload caused by performing additional tasks while driving (Alm, 1993; McLoughlin, Michon, van Winsum, & Webster, 1993; Verwey, 1993). The information overload issue is aggravated by the fact that, until recently, no concepts for the design of a unified and coherent man-machine interface existed. This led Labiale (1997) to remark that juxtaposing these different systems on the dashboard risks making the driving compartment look dangerously like the cockpit of an aircraft.

The classic example of an additional task that impairs driving performance is a mobile phone conversation. While Becker et al. (1995) examined the particularly distracting effects of holding the telephone receiver and dialling a number, Fairclough, Ashby, Ross and Parkes (1990) also measured significant increases in drivers' mental workload when using a hands-free telephone.

2. THE SITUATION-ADAPTIVE MAN-MACHINE INTERFACE

But the hands-free phone will not be the endpoint of technical development. In the early 90s, the DRIVE Project V1041 Generic Intelligent Driver Support GIDS (Michon, 1993) came up with the idea of an adaptive man-machine interface for the optimal allocation of drivers' attentional resources. Verwey (1993) outlines the GIDS MMI concept as a system that schedules information presentation and presents only one message at a time that requires the driver's attention. In addition, message presentation is to be adapted to the current demands exerted on the driver by the driving task. When, for example, a driver is very busy manoeuvring his/her car in heavy city traffic, the interface might adapt to the increase in demands on the driver by blocking phone calls.

Piersma (1991) points out why such a system requires a workload estimator, i.e. a driver model, and why he thinks that workload estimation cannot be based on monitoring physiological measures as implied by Hancock & Chignell (1987). His first argument is that it cannot be assumed that car drivers will accept the use of electrodes in normal driving. Secondly, Piersma argues that measuring the momentary workload supplies no information about the workload in the near future. Thus, determining workload increments by using physiological methods would simply come too late. In order to avoid this, situations resulting in elevated workload have to be predicted before they actually occur. The first argument might be invalidated due to new technical developments in non-obtrusive monitoring methods[1]. Nevertheless, the second argument is still convincing.

Of course, there is controversy over the question of how comprehensive the driver model in the adaptive man-machine system must be. Due to the inherent complexity of the traffic environment and the numerous possible interaction effects with personality traits of the individual driver, Reichart (1993, p. 145) is sceptic about simplistic solutions: “It will be obvious […] that there is still a long way to go before convincing concepts for adaptive interfaces can be introduced”. Verwey (1993) is much more optimistic, maintaining that a simple model consisting of a static look-up table which contains the levels of determining parameters of driver workload may be sufficient. In his view, the crucial point is the identification and classification of relevant parameters that determine workload.

In fact, the workload estimator used in the GIDS project mentioned above was kept very simple. To predict workload, the mean secondary task performance of 24 subjects in six different traffic situations on highways and rural roads was used to estimate the average driver’s spare information processing capacity in these situations (Piersma, 1991; Verwey, 1991). Verwey later extended this technique to examine ten different sorts of traffic situations (Verwey, 2000). In all cases, Verwey used an ad hoc classification system. Apart from the idea of a “small world” made up of a relevant subset of frequently occurring situations (Michon, 1991), there was no systematic approach to the classification of traffic situations. Another problem not addressed in these studies is the question of how oncoming traffic situations can be predicted in real traffic. In contrast to Verwey’s subjects, who followed the instructions of an experimenter, a real driver will not follow a predetermined route.

The GIDS evaluation experiments were either conducted in the form of simulator studies, or were carried out using a provisional system in a test car belonging to the TNO Institute for Perception in Soesterberg, NL. Using this car meant that it was necessary to

[1] This invalidation was pointed out by an anonymous reviewer of a previous draft of this paper.

simulate absent sensor systems, which constrained the number of feasible experimental procedures. For example, it was not possible to measure the distance to leading vehicles; this was only feasible with a specially equipped “stooge” vehicle. The position of the provisional GIDS vehicle had to be tracked with active infrared sensors, which were triggered by reflecting markers along the roadside of a predetermined experimental route (Janssen & Kuiken, 1993). In short, the provisional system, which was installed in a van (and did not have much in common with a normal car), was not particularly realistic.

Nevertheless, GIDS was one of the first experimental systems that implemented adaptivity to the driver (Onken, 1993), and it brought behavioural and technical scientists together in an unprecedented concerted effort to determine the feasibility of in-vehicle knowledge refinery (Michon, 1991). One of the goals of GIDS was to produce a set of recommendations how to avoid confusion and driver overload, while keeping these recommendations quite independent of specific applications (Groeger, Alm, Haller, & Michon, 1993). In fact, there is hardly any problem in the field of driver assistance systems which has not already been addressed in the DRIVE project V1041. But unfortunately, the promising ideas generated ten years ago have not been translated into reality. So we think it’s high time to reconsider the approach of a generic intelligent driver support.

3. WORKLOAD ESTIMATION IN THE CONTEXT OF AN INTEGRATED ASSISTANCE SYSTEM

The goal of advanced driver assistance is still defined as to support safety and comfort by means of an individual, adaptive and integrated human-machine interface (König, Weiß, Gehrke, & Haller, 2000; Mayser, 2002). Integrated assistance means that the individual subsystems are technically linked to each other and operation of the systems is coordinated as an integral unit. An integrated assistance system should also set priorities, depending upon individual needs and situational requirements of the traffic situation. This allows coordination of navigation and communication tasks which overlap in time. The experimental system that we use conforms to these design principles and can be seen as an attempt to renew and advance the GIDS idea by exploiting new technical developments which were not available ten years ago.

The demonstrator vehicle is a modified BMW sedan. It is equipped with the developer version of a state-of-the-art adaptive cruise control system (ACC), which is based on a radar sensor, and an experimental heading control system (HC) based on computer vision. HC searches for lane markings and employs small forces to the steering wheel, which serve as indicators how to steer in order to stay in lane.

Workload estimation is done by a software module which predicts the driver's mental strain and reduces additional mental workload resulting from displays, signals, and system messages by postponing less important messages or cancelling those messages altogether. The module uses input from car sensors and from an experimental digital map which we call an enhanced database for driver assistance systems (EDDAS).

3.1. Pilot Study

As a first step, an experimental route (27 km) in an area covered by the EDDAS map was filmed using a video camera located behind the windscreen of a car. The videotape was used to classify the route according to Fastenmeier's (1995) taxonomy of traffic situations. The beginning of a new situation was defined by a change in any of the six dimensions that constitute the Fastenmeier classification scheme. These are: (1) Road type (5 highway classes, 2 rural road classes, 7 city classes) (2) Horizontal layout (curve versus no curve) (3) Vertical layout (slope versus plane route) (4) Intersections (4 classes) (5) Route constrictions (yes/no) (6) Driving direction (straight ahead, turn left, turn right). 186 consecutive situations could be identified on the route.

In a second step, the geographic coordinates of the transition points from one situation to the next were measured using the means of ten consecutive position fixes collected by a

DGPS[2] receiver. On the basis of this, the Fastenmeier classification of the experimental route was inserted into the EDDAS map.

The 186 situations belong to 22 classes of situations (i.e. 22 different combinations of roadway features). To assess the relative workload induced by each of these classes, we used a secondary task paradigm. Subjects had to perform a visual search task on nine journeys on the experimental route: to look for target words in a text that was slowly scrolled on a display. The task is difficult and would need continuous monitoring to achieve a perfect performance. However, drivers liked this task as we employed acoustic feedback on hits and false alarms, which turned out to be extremely motivating. Subjects had the experience of playing a tricky yet interesting game that they were willing to engage in. As is recommended for the secondary task paradigm (Wickens, 1984, p. 323), we instructed the subjects to consider safe driving as the primary task, and that driving performance should not suffer as a result of the secondary task. To measure the subjects' spare visual capacity, we counted how often the subjects looked at the display. This was done offline by simultaneously split-screen videotaping the driver's face and the traffic environment. According to Rockwell (1988) and Wierwille (1993), glance frequency is a sensitive measure of drivers' visual workload. The number of glances per second devoted to the secondary task display, averaged over all subjects and situations per situation class, varied between .803 (easiest class) and .476 (most demanding class). We concluded that 1 – (mean secondary task glance frequency of situation class) would be a simple yet suitable indicator of drivers' visual workload in the (primary) driving task, which is considered the crucial component of mental workload when driving (Wierda & Aasman, 1992; Fairclough, Ashby, & Parkes, 1993; Wierwille, 1993; Verwey, 2000). This indicator is able to make clear distinctions between traffic situations. In complicated situations, especially at intersections, the secondary task glance frequency is much lower than, for example, on straight rural road segments.

---

[2] DGPS = Differential GPS, a technique that reduces positioning error to ± 3 m.

3.2. System specifications

Workload estimation is carried out in a two-stage process. In stage one, a basic estimate is generated by projectively tracking the EDDAS map to identify the oncoming (statically defined) traffic situations and looking up the respective workload indicators (cf. section 3.1.) for these situations. In stage two, this basic estimate is fine-tuned using information about dynamic aspects of the driving situation.

To be able to track the EDDAS map, the system has to know its exact geographical position. This is achieved using a service based on the global positioning system (GPS) called differential GPS (DGPS). DGPS is a technique used to improve positioning accuracy by determining the positioning error at a known location and subsequently incorporating a corrective factor into the position calculations of DGPS receivers via real-time radio transmission. These DGPS receivers must operate in the same area and simultaneously track the same satellites. Fig. 1 illustrates this principle. Using this technique, positioning error can be reduced to ± 3 m or less.

<< Insert Figure 1 approximately here >>

A map tracking algorithm matches the vehicle position to the EDDAS map and generates a forecast of the route (cf. Schraut, 2000). The accuracy of this projection depends on the level of detail of the digital map and its range depends on the absence of oncoming intersections, as we can never know which route the driver will choose at a junction. However, it is possible to detect whether a driver intends to turn left or right by monitoring the indicator status. Since the driver's route was determined by the experimental procedure of the evaluation experiment (cf. sections 4 to 6), indicator status was not used for route prediction in this experiment.

A look at Fig. 2 should be helpful to understand how the system actually works. The upper half shows a viewer application used to monitor operation of the workload estimator. The lower half of Fig. 2 is the plot of a typical 6 minute-long section of a workload estimation record. The top pane in the viewer's screenshot shows a symbolic representation of the route projection. The bullets identify situations.

<< Insert Figure 2 approximately here >>

The workload indicator for each situation class is read from a look-up table, thus transforming the sequence of situations into a sequence of workload level projections. The time at which a given point or situation will be reached is anticipated on the basis of the current driving speed. The greater the distance to a point, the more uncertainty lies in this forecast in the time domain, since the assumption that driving speed will remain constant is a coarse simplification. Nevertheless, the fast iteration cycle (10 Hz) of the computations makes the forecast precise enough for situations in the vicinity. One question here is how many situations ahead should be considered, or in other terms, how far an average driver could be expected to look ahead.

Verwey's (2000) guess is that intersections and roundabouts should be assumed to "begin" 4 s before the situation is actually encountered, in order to take into account the workload caused while approaching these road situations. While this seems plausible, there is no theoretical foundation for this value. We wanted to use the maximum look-ahead distance actually used by drivers. It is known that drivers do lower their driving speeds on motorways on account of reduced visibility conditions in fog when visibility is about 300 m or lower (Hawkins, 1988), so 300 m should be a good estimate. Look-ahead is most relevant when driving at high speeds. On our test route, which does not include highway driving, the maximum legally allowed speed is 100 km/h. At a speed of 100 km/h, a 300 m look-ahead corresponds to 10.8 s time headway, but it is not uncommon and is legally allowed to drive at speeds of 200 km/h and more on German motorways, which reduces the 300 m look-ahead to just about 5 s and less in the time domain. On the basis of these considerations, we derived our own proposal for generic proximity weighting of the workload forecast, which can be seen in Fig. 2. The upcoming 5 s are taken into account fully, and then weighting is applied with an exponential first order decay ($y = 2.71866\ e^{-x/4.72657}$). It should be noted that the decay function is only an "educated guess", inspired by the idea that the importance of remote situations is diminishing. The time-integrated value of the weighted workload forecasts is used as the basic workload estimate.

This basic workload estimate, which dependends only on the situations ahead, is refined using knowledge about dynamic aspects of the traffic environment and variables of vehicle dynamics which are indicative of critical driving manoeuvres. Fig. 3 shows these refinements; notation is simple: $w = 0.92$ means that the workload estimation value is set to 0.92 and $w = 1.1 * w$ means that w after this weighting step is the old value of w multiplied by 1.1. Most steps are multiplications with a weighting factor, only an ACC request to the driver to take over control results in an assignment of a constant, overriding previous computations. The reason for this will become clear later, when we explain why these ACC requests occur.

We distinguish between free driving and following a leading car. As Weinberger (2001) explicates, drivers of ACC vehicles often feel uneasy when they are approaching a slower leading vehicle. This is because in contrast to some early prototypes, today's adaptive cruise control systems do not automatically handle every slow-down manoeuvre necessitated by a slower leading vehicle. The maximum deceleration the ACC can request from the brake system is typically limited to about $-2\ m/s^2$. This leads to situations where the driver has to take over control by pressing the brake pedal (pressing the brake pedal deactivates the ACC). In some situations, it is not perfectly clear whether taking over control will be needed or not, so an ACC-equipped vehicle may require some additional attention from the driver when approaching a leading vehicle. Moreover, drivers do wait to see if the ACC will cope with the situation and then react later by decelerating rapidly if the system does not recover the time gap sufficiently (Marsden, McDonald & Brackstone, 2001).

In the course of some informal tests, subjects told us that compared to free driving on an empty road, they subjectively felt a slight increase in workload when following a leading vehicle. However, the subjects felt a considerable increase in workload whenever a close following distance decreased even further. Accordingly, the mere fact of detecting a leading vehicle (the radar sensor has a maximum "viewing distance" of 120 m) is taken into account by multiplying the basic workload estimate by a factor of 1.1. However, when rapid approaches are detected at a point when time headway amounts to 3 s or less, the workload estimate is incremented considerably at once, as can be seen from Fig. 3. The shortest time headways we measured were approximately 0.5 s. The system should not be made too hesitant in responding to rapid approaches, because detection of a close target can occur abruptly, especially in very narrow curves, where the radar system can run into problems because of its limited "field of view" (8° horizontally).

The distinction we make between free driving, following and rapid approach fits well into the three-phase concept of approaching a leading vehicle proposed by Vogel (2002). In this concept, a free vehicle at first becomes an influenced vehicle at a 6 s time headway and then a following vehicle at a 2-3 s time headway. Based on the analysis of 110,269 headways of following vehicles, Vogel argues that drivers in fact do drive at time headways of between 2 and 3 s, as the correlation between the speed of the leading and the following vehicle lies between approximately .50 and .70 in this time headway interval. But correlations already begin to rise at a 6 s time headway, while they are consistently low at headways greater than 6 s. The threshold of 6 s can therefore be seen as the line of demarcation between driving freely and driving under some influence of the vehicle ahead. This time headway of 6 s corresponds to 83 m at 50 km/h and to 167 m at 100 km/h. It is still unclear whether the 6 s threshold is psychologically salient (Vogel, 2002). Moreover, the 6 s threshold was derived from measurements at an urban junction, so its validity may be restricted to this traffic situation. Therefore, we consider it wise to use the information for a detected leading vehicle at a distance headway of up to 120 m (i.e. the maximum distance detected by the ACC radar) to moderately increase the workload estimate, as the possibility of a slightly premature increase in estimation of the workload is more acceptable than a delayed one.

If a software detector encounters an intersection with traffic lights on the route projected in the time domain within the next 5 s, both arms of the fine tuning logic apply a factor of 1.1, as an approach to such an intersection is considered particularly complicated (Fastenmeier, 1995). The 5 s look-ahead applied here is a slight extension of Verwey's (2000) proposal of 4 s as mentioned above.

<< Insert Figure 3 approximately here >>

If no leading vehicle is detected but the driver hits the brakes hard, we also add a considerable increment to the workload estimate, as can be seen from Fig. 3. Braking hard is defined as a deceleration of more than -1 m/s^2. Here, we apply a less extreme increment than that used when a rapid approach is detected by ACC sensor data, as rather hard decelerations may just be the result of a sporty driving style. Nevertheless, the rapid approach and hard braking detectors address similar situations. The approach detector simply captures some situations in advance where the driver may need to brake very soon. While accelerations are never critical manoeuvres in volume production vehicles, extreme decelerations are indicative of situations that induce workload. This can be illustrated by the fact that Weinberger (2001) finds mean decelerations of more than -2 m/s^2 in all his critical situations, during which drivers took control and overrode the ACC system.

By ACC take-over request (called "committal" in Fig. 3 for brevity) we mean that ACC commits control back to the driver. In the demonstrator vehicle, which is used to evaluate the system, optical and acoustical signals prompt the driver to take over the longitudinal vehicle control whenever the ACC-induced deceleration is not sufficient to avoid a crash or when other restrictions apply. For example, the ACC switches itself off if the vehicle's speed becomes slower than 30 km/h, a situation that typically arises when approaching red traffic lights. The acoustic prompt is useful to beginners, but annoying to experienced ACC drivers, who prefer to switch off the auditory signal. Whether experienced or not, the exact point in time when the driver has to take over control always signifies mental strain. On average, subjects judged the amount of "activation" they felt in these moments to be 92% of an imaginary maximum. As ACC committal (take-over request) is a very distinct occurrence that always induces workload, we set the workload estimate to .92 for the time at which the prompting signals are displayed.

Finally, we must explain the overtaking detector (cf. Fig. 3). Overtaking is defined as a situation when the left indicator is in operation at driving speeds faster than 70 km/h or at points on the route where there are no intersections (including T-junctions) in the route projection while applying a look-ahead of 10 s. As this detector is defined in the left arm of the fine tuning logic, this conditional weighting begins at the very moment when the radar sensor loses "sight" of the leading vehicle, i.e. when overtaking is actually in progress. The factor of 1.44 applied to the workload estimate in these moments corresponds to the reduction in secondary task glances whilst overtaking that we found during the analysis of overtaking situations from the pilot study (cf. section 3.1).

Whenever the finely tuned workload estimate exceeds a threshold value, incoming telephone calls are redirected to the mailbox, instead of signalling them to the driver. This is exactly what we carried out for the purpose of evaluation in the demonstrator car. Of course, it is possible to "silence" short messages, email, and additional displays of all kind in the same way.

The threshold value of .35 depicted as a straight line in the lower half of Fig. 2 is the value chosen for the evaluation experiment, during which it was necessary to "sensitise" the workload estimator to a sufficient degree in order to be able to suppress a considerable amount of incoming telephone calls on our test route. A somewhat higher threshold would have to be applied in a real-life situation.

4. METHOD

The method choosen for a preliminary evaluation study of the system was a field experiment. Participants consisted of 6 younger male drivers who were from 21 to 29 years old (M: 24.28 years, SD: 3.60 years), and 6 male novice drivers (M: 18.26 years, SD: 0.33 years). The experienced drivers were holding a driving license for at least some years (M: 6.80 years, SD: 3.29 years), novices had gotten their driving licence 50 days before the experiment on average (M: 0.14 years, SD: 0.08 years). Experienced drivers were recruited from the student population of Universität Regensburg, novices from a local driving school. All subjects were paid for their participation.

Each subject drove the experimental route 3 times; the sequence of experimental conditions was permuted within groups. None of the subjects was familiar with the experimental route. The experimenter provided standardized route guidance instructions throughout the whole experiment. The experimental conditions (within subjects variable) were: driving without driver assistance, driving with ACC and HC, and driving with ACC, HC, and adaptive telephone behaviour (called ATB in the following), which was controlled by workload estimation.

Incoming telephone calls were only blocked in the workload estimation condition (ATB) when the workload estimate exceeded the threshold value. In all conditions, the driver was required to respond to ten mental arithmetic tasks via a hands-free telephone. Subjects answered the phone by pressing a button integrated in the steering wheel. The arithmetic tasks consisted of adding 12 to a number in the range of 11 to 93 and saying the answer aloud. This phone task simulated a short real-life telephone conversation with a medium level of complexity. Phone calls were placed from a second experimenter (caller) on a mobile phone and were spaced roughly 2 minutes apart. In the workload estimation condition (ATB), the caller required up to 64 attempts to make a call, as the workload estimator rejected up to 54 incoming calls. To determine whether the assistance systems have a beneficial effect on driver workload, the mental effort was measured for all three experimental runs using objective (electrocardiogram ECG, electromyogram EMG) and subjective measures (offline ratings from observers watching video scenes, NASA TLX self-report scale).

Recording of physiological data was done with a Becker Meditec Varioport data logger, the ECG was sampled at a rate of 256 Hz and the integrated EMG was sampled at a rate of 4 Hz.

EGC was measured with a lead from the manubrium sterni to the lowest left rib, this chest lead was used because it minimizes movement artefacts. R-spike detection was done offline with a slightly modified version[3] of the EP Limited (2000) Open Source implementation of the Pan-Tompkins detector (Pan & Tompkins, 1985). All ECG data were visually inspected, because it is considered bad practice to rely on automatic artefact detection methods (Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996, p. 1051). Artefact correction was done according to the guidelines of Mulder (1988, 1992). Heart rate (HR) in beats per minute was used as the most simple indicator of drivers' workload (Hering, 1999). It is known that an increase in HR primarily measures physical effort, although it does not discriminate between mental effort and physical effort (Manzey, 1998).

For this reason, heart rate variability (HRV), defined as the integrated power spectral density (PSD) in the frequency band 0.07 Hz to 0.14 Hz was also computed. A decrease in HRV is known to indicate an increase in the mental effort invested by the subject (Mulder, 1988, 1992). Spectral analysis of the ECG data was done with a standard implementation of the FFT (cf. Press, Teukolsky, Vetterling & Flannery, 1992, chap. 13). To generate an equidistant time series, the interval function was interpolated linearly and resamled at 8 Hz. This results in a time series which is known to be equivalent to the interval function (Mulder,

[3] Implementation details of the software written for ECG analysis are beyond the scope of this paper. This software, including all C++ sources, can be requested free of charge from the corresponding author.

1992) and apt for a standard FFT. Expert groups (Task Force, 1996; Berntson et al., 1997) still prefer this method over alternative approaches of spectral analysis. After a Hann-windowing of the equidistant time series, zero padding to the next power of 2 was done, and then the FFT was computed.

A third physiological measure was the lateral frontalis electromyogram (EMG) as defined in the guidelines of Fridlund & Cacioppo (1986). This EMG lead is named after the musculus frontalis, where the electrodes are applied, but it actually is a combined record of m. frontalis' as well as other face muscles' activation. In contemporary psychophysiological EMG research, face muscle leads are preferred over the formerly used skeleton muscle leads (Manzey, 1998), and the frontalis lead is supposedly most diagnostic for measuring mental effort (de Waard, 1996). The EMG level rises when exertion increases. It was defined as the arithmetic mean of the 4 Hz samples of integrated EMG in this study.

Observer ratings were also used as an additional approach to assess the workload induced by a given phone call. All the 360 telephone conversations (mental arithmetic tasks) had been videotaped with a split-screen recording from 4 different camera angles, with one of the 4 channels showing the traffic scene in front of the car, another showing the driver's face. We asked 20 subjects to judge all of the telephone scenes. None of these subjects had any experience with the experimental vehicle, nor did they know that the scenes had been recorded under different experimental conditions. They were asked to rate how disturbing each individual telephone call was on a six-point scale.

The 12 drivers of the field experiment also responded to NASA TLX self-report scales for mental workload after each ride. We used the TLX version adapted to driving by Fairclough (1991) and computed the so-called raw TLX (RTLX), an unweighted sum of the subscale values. In the original version of the TLX (Hart & Staveland, 1988), paired comparisons were used to derive weights for the six subscales of the TLX. But, as Byers, Bittner, & Hill (1989) already showed, RTLX scores can even give a better account of the workload experienced by the subject than traditional TLX values. Since Fairclough's RTLX version uses a modified subscale definition, the results collected with these scales must not be mistakenly compared with other versions that use the common TLX scale definitions of Hart and Staveland.

5. RESULTS

An alpha level of .05 was used for all statistical tests. A repeated measures ANOVA of the RTLX sum scores revealed only a marginally significant difference between the experience groups ($F(1,10) = 4.391$, $p = .0626$), but no other significant effects. Beginners felt a higher amount of mental strain than experienced drivers, regardless of the experimental condition.

The reliability of the video observer ratings was tested using an intraclass correlation coefficient (McGraw & Wong, 1996). ICC (A,20) = .924 indicates a high level of absolute agreement between raters. In the case of experienced drivers, the telephone conversations "allowed" by the workload estimator in the adaptive telephone behaviour (ATB) condition were rated as less disturbing than in the base condition (Wilcoxon $z = -4.586$, $p < 0.001$) and also as less disturbing than in the ride with heading control and adaptive cruise control (Wilcoxon $z = -3.608$, $p < 0.001$). This effect was not ascertained in the case of the novices (Fig. 4, lower right).

All objective workload measures were subjected to a 2×3 repeated measures ANOVA (experience × condition). Driving experience was a between subjects factor and experimental condition a within subjects factor. Unfortunetely, neither main effects nor the interaction were significant for all measures. Even so, we think the tendencies found are noteworthy, as all the results point in the same direction (cf. Fig. 4).

<< Insert Figure 4 approximately here >>

Heart Rate variability seems elevated in the adaptive telephone behaviour condition (denoted HC+ACC+ATB in Fig. 4, upper left) compared to driving under the other conditions. This would indicate a reduction in mental workload, but as said above, the effect is not significant ($F(2,9) = 1.043$, $p > .39$). Also, heart rate seems to be reduced in the condition with adaptive telephone behavior (Fig. 4, upper right), but again the effect is not significant ($F(2,9) = 0.922$, $p > .43$). Visual inspection suggests somewhat lower tonic muscle tension in the case of experienced drivers (Fig. 4, lower left), but there is no significant main effect of driving condition ($F(2,9) = 1.0381$, $p > .39$) nor an interaction ($F(2,9) = 1.9565$, $p > .20$).

6. DISCUSSION

Today, estimating real-time workload is not much more than a synthesis of heuristics and rules of thumb, as there is still only sparse knowledge about the precise impact of different situation-specific influences on drivers' mental workload. Studies that deal with measuring drivers' mental workload such as those carried out by de Waard (1996) or Verwey & Veltman (1996) concentrate on offline measurements that are not directly applicable in real-time workload estimation. Clearly, Verwey (2000) and, to some extent, also de Waard (1996), foresee the indirect method of detecting classes of situations with known mean effects on mental workload from an onboard geographical database. However, the lack of a unified classification scheme for traffic situations severely hinders comparisons between results of different studies found in the literature. One candidate for a common vocabulary for describing traffic situations is Fastenmeier's (1995) taxonomy, which has already received acceptance in German psychology, ergonomics, and road construction. Therefore, it was used in this study.

All in all, the results of the preliminary system evaluation show a reduction of mental workload in the adaptive telephone condition, where incoming telephone calls are not signalled to the driver, but redirected to the mailbox whenever the workload estimation exceeds threshold. Video ratings tell us that while the workload for experienced drivers is reduced significantly, this is not the case for beginners. For possible reasons, see below.

There are only tendencies in the psychophysiological measurements that indicate a workload reduction for both experienced and novice drivers. While this seems disappointing, it also is clear that a sample of just 12 subjects has to be considered tiny for psychophysiological methods. Hering (1999) used the heart rate data of 100 drivers in his study of situational influences on heart rate. Unfortunately, it seems that such expensive experiments may also be needed if one wants to base the evaluation of adaptive driver assistance systems on psychophysiological measurements, which have always been appealing to many researchers on account of their apparent objectivity.

However, there are no statistically significant effects of reduced workload for the group of novice drivers. Some observations of clearly reduced driving performance in the beginners group during the telephone conversations (e.g. driving very close to the curb or suddenly turning left at an intersection which should be driven straight through) suggest that they suffered from a lack of situational awareness, i.e. novices seem to allocate too many cognitive resources to the phone task, regardless of how complicated the traffic situation. If novices are simply unaware of critical situations that require divided attention in order to be able to drive safely, e.g. while being engaged in a telephone conversation, then it is not surprising that we were unable to detect any workload reduction with adaptive telephone behaviour. This would mean that it takes some driving experience in the first place to learn how to even detect critical situations and hence become aware that these situations require additional mental effort. Only further research can tell us if this interpretation is true. Some corroboration of this hypothesis can be found in observation of gaze behavior: Novice drivers do not change their scanning patterns as situations become more complex, but experienced drivers do (Underwood, Chapman, Bowden, & Crundall, 2002).

When judging what was really achieved in these experiments, it is important to understand and appreciate that the goal of the reported work was to explore and demonstrate new possibilities of creating a “situation-aware” vehicle by means of a synthesis of human factors knowledge and state-of-the-art technology. The demonstrator vehicle turned out to be a stable research platform that permits field tests in everyday traffic. Hence, the results have high ecological validity. While there is still an overwhelming amount of unresolved questions that require in-depth analysis, the workload estimator presented here may be at least one further step towards producing real-time estimates of workload for drivers in real traffic. Situation-adaptive automotive applications could become a dream come true in the near future if the methods and technology in this area are developed further. An additional necessary prerequisite for this vision is a seamless integration of the new technologies or devices into the man-machine interface of the vehicle. We feel that we no longer need to “appreciate the archetypical significance of fairy tales” (Michon, 1991, p. 13) when speculating about situation-adaptive man-machine interfaces in future vehicles. These intelligent systems are already in reach and have the potential to enhance comfort and safety of the driving task.

ACKNOWLEDGEMENTS

We thank Rainer Scheuchenpflug and two anonymous reviewers for helpful comments on earlier drafts of the manuscript. The study was a part of the project SANTOS (Situation-adapted and user-type optimized driver assistance systems), which was supported by the German Federal Ministry of Education and Research (BMBF) under grants no. 19S9826A/B. The authors bear responsibility for the content of this paper.

REFERENCES

Alm, H. (1993). Route navigation. Deciding driving information needs. In A.M. Parkes & S. Franzen (Ed.), Driving future vehicles (pp. 187-192). London: Taylor & Francis.

Becker, S., Brockmann, M., Bruckmayr, E., Hofmann, O., Krause, R., Mertens, A., Nin, R., & Sonntag, J. (1995). Telefonieren am Steuer. Berichte der Bundesanstalt für Straßenwesen. Mensch und Sicherheit. Heft M 45. Bremerhaven: Wirtschaftsverlag NW.

Berntson, G.G., Bigger, J.T., Eckberg, D.L., Grossman, P.G., Kaufmann, P.G., Malik, M., Nagaraja, H.N., Porges, S.W., Saul, J.P., Stone, P.H., & Van der Molen, M.W. (1997).

Heart rate variability: Origins, methods, and interpretive caveats. Psychophysiology, 34, 623-648.

Byers, J.C., Bittner, A.C., & Hill, S.G. (1989). Traditional and raw task load index (TLX) correlations: Are paired comparisons necessary? In A. Mital, J. Fernandez, H. Krueger, S. Kumar, & M. Menozzi (Eds.) Advances in Occupational Ergonomics & Safety I : Proceedings of the XIth Annual International Occupational Ergonomics and Safety Conference Held in Zurich, Switzerland, 8-11 July 1996, (pp. 481-485). London: Taylor & Francis.

De Waard, D. (1996). The measurement of drivers' mental workload. Haren: The Traffic Research Center VSC, University of Groningen.

EP Limited (2000). Open Source arrhythmia detection software. [On-line]. Available: http://www.eplimited.com/software.htm

Fairclough, S.H., Ashby, M.C., Ross, T., & Parkes, A.M. (1990). Effects on driving behaviour of handsfree telephone use. DRIVE Project V1017 (BERTIE) Report No. 48. Loughborough: HUSAT Research Institute, November 1990.

Fairclough, S.H. (1991). Adapting the TLX to measure driver mental workload. DRIVE Project V1017 (BERTIE) Report No. 71. Loughborough: HUSAT Research Institute, November 1991.

Fairclough, S.H., Ashby, M.C., & Parkes, A.M. (1993). In-vehicle displays, visual workload and usability evaluation. In A.G. Gale et al. (Eds.), Vision in vehicles IV (pp. 245-254). Amsterdam: Elsevier Science Publishers B.V.

Fastenmeier, W. (1995). Die Verkehrssituation als Analyseeinheit im Verkehrssystem. In W. Fastenmeier (Ed.), Autofahrer und Verkehrssituation: neue Wege zur Bewertung von Sicherheit und Zuverlässigkeit moderner Sraßenverkehrssysteme (pp. 27-78). Köln: Verl. TÜV Rheinland.

Fridlund, A.J., & Cacioppo, J.T. (1986). Guidelines for human electromyographic research. Psychophysiology, 23, 567-589.

Groeger, J.A., Alm, H. Haller, R., & Michon, J.A. (1993). Impact and acceptance. In J.A. Michon (Ed.), Generic intelligent driver support (pp. 217-239). London: Taylor & Francis

Hancock, P.A., & Chignell, M.H. (1987). Adaptive control in human-machine systems. In P.A. Hancock (Ed.), Human factors psychology (pp. 305-345). Amsterdam: Elsevier.

Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results and theoretical research. In P. Hancock & N. Meshkati (Eds.). Human mental workload (pp. 139-183). Amsterdam: North Holland.

Hawkins, R.K. (1988). Motorway traffic behaviour in reduced visibility conditions. In A.G. Gale et al. (Eds.), Vision in vehicles II (pp. 9-18). Amsterdam: Elsevier.

Hering, K. (1999). Situationsabhängiges Verfahren zur standardisierten Messung der kognitiven Beanspruchung im Straßenverkehr. Aachen: Shaker.

Janssen, W.H., & Kuiken, M.J. (1993). Evaluation studies. In A.J. Michon (Ed.), Generic intelligent driver support (pp. 195-215). London: Taylor & Francis.

König, W., Weiß, K.E., Gehrke, H., & Haller, R. (2000). S.A.N.T.O.S Situationsangepasste und Nutzer-Typ-zentrierte Optimierung von Systemen zur Fahrerunterstützung. In H. Bubb (Ed.), Ergonomie und Verkehrssicherheit. Dokumentation der Herbstkonferenz der GfA. München, 12.-13.10.2000 (pp. 107-113). München: Utz.

Labiale, G. (1997). Cognitive ergonomics and intelligent systems in the automobile. In Y.I. Noy (Ed.), Ergonomics and safety of intelligent driver interfaces (pp. 169-184). Mahwah, NJ: Lawrence Erlbaum.

Manzey, D. (1998). Psychophysiologie mentaler Beanspruchung. In F. Rösler (Ed.), Ergebnisse und Anwendungen der Psychophysiologie. Enzyklopädie der Psychologie Vol. C/I/5 (pp. 799-864). Göttingen: Hogrefe.

Marsden, G., McDonald, M., & Brackstone, M. (2001). Towards an understanding of adaptive cruise control. Transportation Research Part C, 9, 33-51.

Mayser, C. (2002). An advanced concept for integrated driver assistance systems. In Proceedings of the 9th World Congress on Intelligent Transport Systems in Chicago, October 14-17, 2002.

McGraw, K.O., & Wong, S.P. (1996). Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1, 30-46.

McLoughlin, H.B., Michon, J.A., van Winsum, W., & Webster, E. (1993). GIDS intelligence. In A.J. Michon (Ed.), Generic intelligent driver support (pp. 89-112). London: Taylor & Francis.

Michon, J.A. (1991). A vision of the future. In A.G. Gale et al. (Eds.), Vision in vehicles III (pp. 3-15). Amsterdam: Elsevier.

Michon, J.A. (Ed.) (1993). Generic intelligent driver support. A comprehensive report on GIDS. London: Taylor & Francis.

Mulder, L. J. M. (1988). Assessment of cardiovascular reactivity by means of spectral analysis. Ph.D. Thesis, Groningen University Press, University of Groningen, NL.

Mulder, L. J. M. (1992). Measurement and analysis methods of heart rate and respiration for use in applied environments. Biological Psychology, 34, 205-236.

Onken, R. (1993). What should the vehicle know about the driver? In A.M. Parkes & S. Franzen (Eds.), Driving future vehicles (pp. 49-67). London: Taylor & Francis.

Pan, J. & Tompkins, W. J. (1985). A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering, 32, 230-236.

Piersma, E.H. (1991). Real time modelling of user workload. In Y. Quéinnec & F. Daniellou (Eds.), Designing for everyone (pp. 1547-1549). London: Taylor & Francis.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P. (1992). Numerical recipes in C: the art of scientific computing. (2nd ed.). Cambridge University Press.

Reichart, G. (1993). Problems in vehicle systems. In A.M. Parkes & S. Franzen (Eds.), Driving future vehicles (pp. 143-146). London: Taylor & Francis.

Rockwell, T. H. (1988). Spare visual capacity in driving—revisited: new empirical results for an old idea. In A.G. Gale et al. (Eds.), Vision in vehicles II (pp. 317-324) Amsterdam: Elsevier.

Schraut, M.P. (2000). Umgebungserfassung auf Basis lernender digitaler Karten zur vorausschauenden Konditionierung von Fahrerassistenzsystemen. Doctoral thesis. On-line available: http://tumb1.biblio.tu-muenchen.de/publ/diss/ei/2000/schraut.html

Sprenger, A. (1999). Ablenkung durch MMI-Lösungen. Der Blick zu Instrumenten im Fahrzeug als Unfallursache. In F. Meyer-Gramcko (Ed.), Verkehrspsychologie auf neuen Wegen: Herausforderungen von Strasse, Wasser, Luft und Schiene (pp. 148-156). Bonn: Dt. Psychologen-Verl.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. Circulation, 93, 1043-1065.

Underwood, G., Chapman, P., Bowden, K., & Crundall, D. (2002). Visual search while driving: skill and awareness during inspection of the scene. Transportation Research Part F, 5, 87-97.

Verwey, W.B. (1991). Adaptive interfaces based on driver resource demands. In Y. Quéinnec & F. Daniellou (Eds.), Designing for everyone (pp. 1541-1543). London: Taylor & Francis.

Verwey, W.B. (1993). How can we prevent overload of the driver?, In A.M. Parkes & S. Franzen (Eds.), Driving future vehicles (pp. 235-244). London: Taylor & Francis.

Verwey, W.B. (2000). On-line driver workload estimation. Effects of road situation and age on secondary task measures. Ergonomics, 43, 187-209.

Verwey, W.B. & Veltman, J.A. (1996). Detecting short periods of elevated workload. A comparison of nine common workload assessment techniques. Journal of Experimental Psychology: Applied, 2, 270-285.

Vogel, K. (2002). What characterizes a „free vehicle" in an urban area? Transportation Research Part F, 5, 313-327.

Weinberger, M. (2001). Der Einfluß von Adaptive Cruise Control auf das Fahrerverhalten. Aachen: Shaker.

Wickens, C.D. (1984). Engineering psychology and human performance. Columbus, OH: Merrill.

Wierda, M. & Aasmann, J. (1992). Seeing and driving: computation, algorithms and implementation. Traffic Research Centre, University of Groningen, The Netherlands.

Wierwille, W.W. (1993). An initial model of visual sampling of in-car displays and controls. In A.G. Gale et al. (Eds.), Vision in vehicles IV (pp. 271-280). Amsterdam: Elsevier.

FIGURE CAPTIONS

Fig. 1: Principle of route projection. The car's position on route is determined to ± 3 m by means of differential GPS technology. A detailed digital map is used to predict upcoming situations and anticipate driver workload.

Fig. 2. Screenshot of running workload estimation system and 6 minutes of an estimation time series (3600 estimates).

Fig. 3. Fine tuning of the workload estimate (see text for details).

Fig. 4 Physiological measures (95 % confidence intervals) and mean observer ratings (lower right). Experimental conditions are driving without any assistance (base), driving with heading control and adaptive cruise control (HC+ACC), and driving with HC, ACC and adaptive telephone behaviour (ATB).

Figure 1

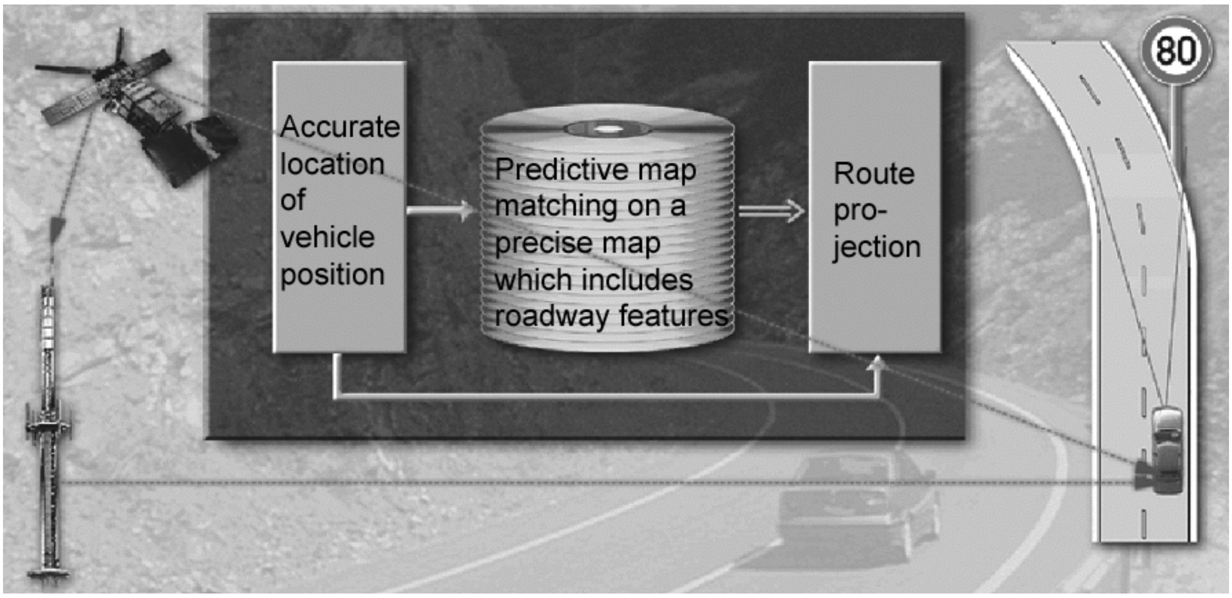
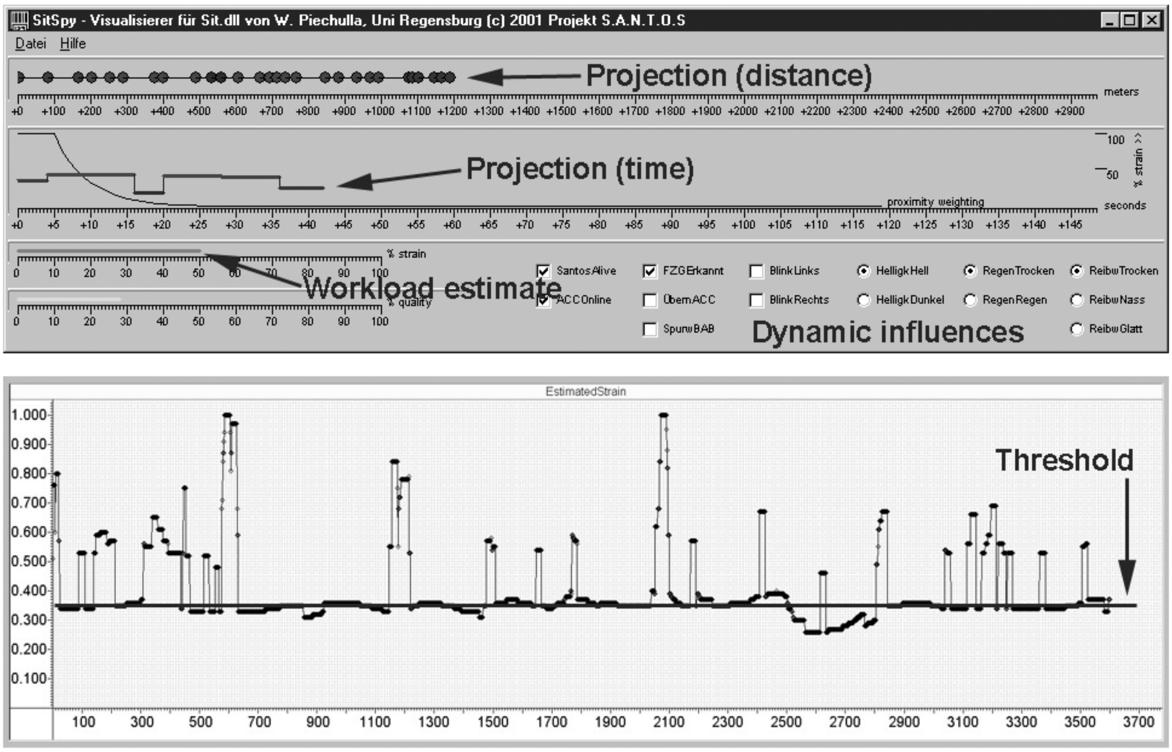

Accurate location of vehicle position
Predictive map matching on a precise map which includes roadway features
Route pro- jection
80

Figure 2



SitSpy - Visualisierer für Sit.dll von W. Piechulla, Uni Regensburg (c) 2001 Projekt S.A.N.T.O.S
Datei
Hilfe
Projection (distance)
meters
Projection (time)
proximity weighting
seconds
% strain
Workload estimate
% quality
SantosAlive
ACCOnline
FZGErkannt
ÜbernACC
SpurwBAB
BlinkLinks
BlinkRechts
HelligkHell
HelligkDunkel
RegenTrocken
RegenRegen
ReibwTrocken
ReibwNass
ReibwGlatt
Dynamic influences
EstimatedStrain
Threshold

Figure 3

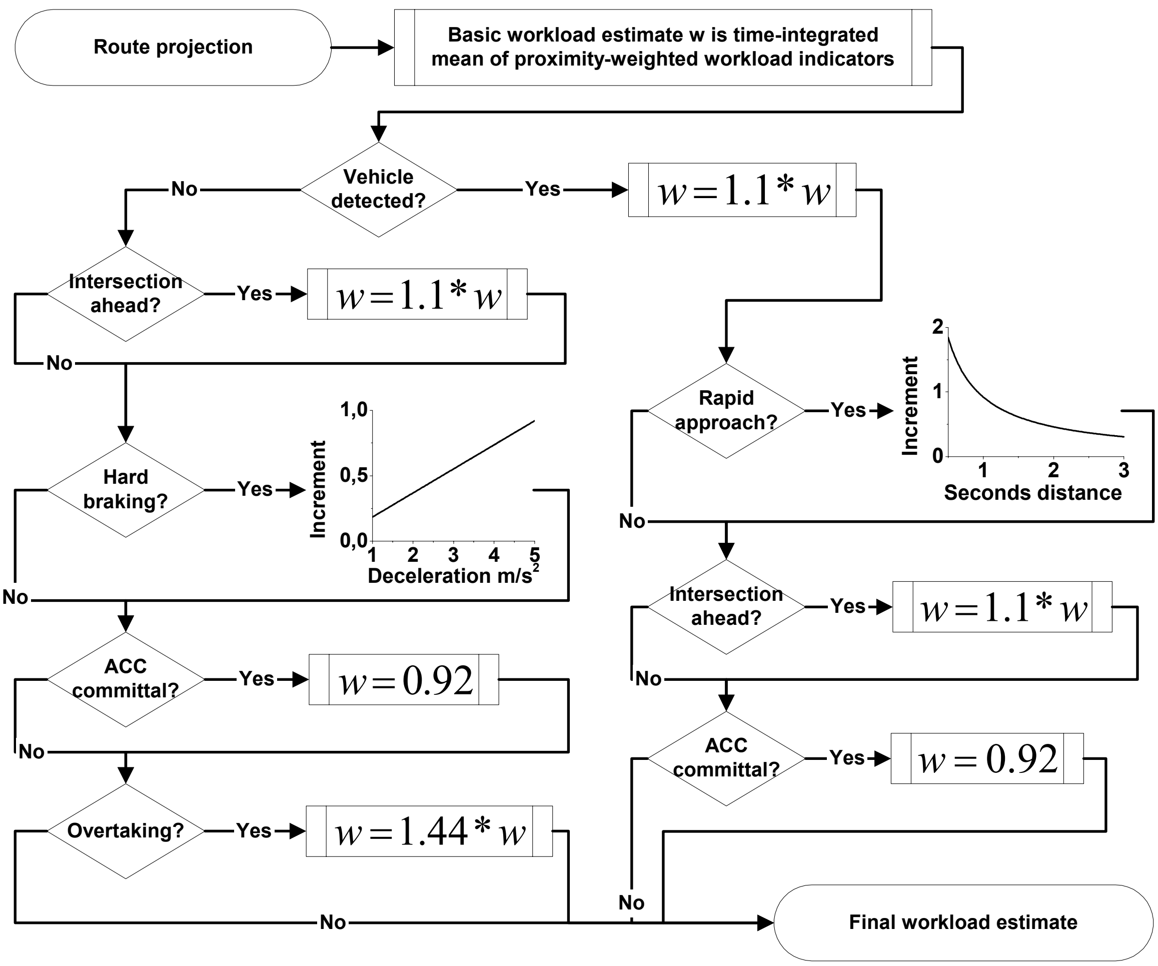
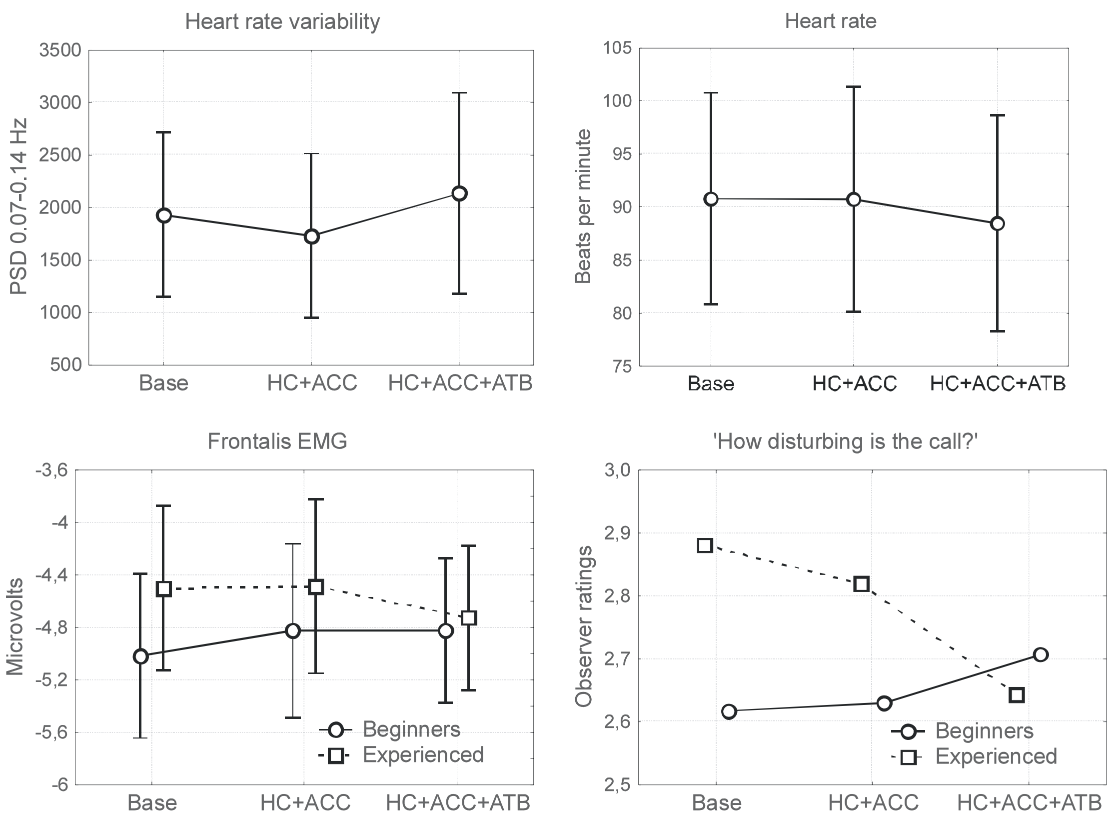

Route projection
Basic workload estimate w is time-integrated mean of proximity-weighted workload indicators
Vehicle detected?
Yes
No
w = 1.1 * w
Intersection ahead?
Yes
No
w = 1.1 * w
Hard braking?
Yes
No
Increment
1,0
0,5
0,0
1 2 3 4 5
Deceleration m/s²
ACC committal?
Yes
No
w = 0.92
Overtaking?
Yes
No
w = 1.44 * w
Rapid approach?
Yes
No
Increment
2
1
0
1 2 3
Seconds distance
Intersection ahead?
Yes
No
w = 1.1 * w
ACC committal?
Yes
No
w = 0.92
Final workload estimate

Figure 4



Heart rate variability
PSD 0.07-0.14 Hz
3500
3000
2500
2000
1500
1000
500
Base
HC+ACC
HC+ACC+ATB
Heart rate
Beats per minute
105
100
95
90
85
80
75
Base
HC+ACC
HC+ACC+ATB
Frontalis EMG
Microvolts
-3,6
-4
-4,4
-4,8
-5,2
-5,6
-6
Beginners
Experienced
Base
HC+ACC
HC+ACC+ATB
'How disturbing is the call?'
Observer ratings
3,0
2,9
2,8
2,7
2,6
2,5
Beginners
Experienced
Base
HC+ACC
HC+ACC+ATB