Applied Ergonomics xxx (2009) 1–7

Contents lists available at ScienceDirect

Applied Ergonomics

journal homepage: www.elsevier.com/locate/apergo

Time perception as a workload measure in simulated car driving

Daniel Baldauf[a,b,*], Esther Burgard[a], Marc Wittmann[a,c]

[a] Generation Research Program, Human Science Center, Ludwig-Maximilians-University of Munich, Prof.-Max-Lange-Platz 11, 83646 Bad Tölz, Germany
[b] Department of Psychology, Ludwig-Maximilians-University Leopoldstasse 13, 80802 Munich, Germany
[c] Department of Psychiatry, University of California San Diego, 9500 Gilman Drive, CA 92093 – 9116A, La Jolla, USA

ARTICLE INFO

Article history:
Received 8 July 2007
Accepted 8 January 2009

Keywords:
Mental workload
Time perception
Duration production
Driving simulator

ABSTRACT

In experimental studies using flight simulations subjects' duration estimates have shown to be an effective indicator of cognitive task demands. In this study we wanted to find out whether subjective time perception could serve as a measure of cognitive workload during simulated car driving. Participants drove on a round course of a driving simulator consisting of three different environments with different levels of task demands. Drivers were required to perform a time-production task while driving the vehicle. Electrodermal activity and subjective ratings of mental workload (SWAT) were recorded simultaneously. The length of produced intervals increased significantly in more complex driving situations, as did electrodermal activity and subjective ratings of mental workload. Thus, time production is a valid indicator of cognitive involvement in simulated driving and could become a valid method to measure the current mental workload of car drivers in various traffic situations.

1. Introduction

When we steer a car and maneuver through various traffic situations, high demands are placed on our central-processing resources. In this context, mental workload is defined as the overall cognitive effort a person invests in his performance while carrying out a task (Hart and Wickens, 1990). As this definition implies that it is a complex, multidimensional phenomenon, various measurements have to be employed to assess the mental workload involved in operating machines and devices (Gopher and Donchin, 1986; Hancock and Mechkati, 1988; Bao et al., 2002). Besides primary task performance, subjective rating scales and physiological parameters are the most widely used measures (Wittmann et al., 2006). To effectively measure secondary task performance as an indicator of cognitive workload, two prerequisites have to be fulfilled: 1) secondary task performance must not interfere with the performance in the primary task (see Wierwille and Connor, 1983; Wickens, 1992), and 2) secondary tasks have to directly compete with the primary task for the same resources. According to Jahn et al. (2003), a major advantage of secondary task measurements is the possibility to distinguish between subtle differences in workload demands that do not affect the primary task performance. The workload demands of the primary task, e.g., driving a car, can then be inferred from the secondary task performance since primary and secondary task performances are supposed to be inversely proportional to each other (Johansson et al., 2004). In the context of driving, two secondary tasks are frequently used: the *paced auditory serial test* (PASAT) and the *peripheral detection task* (PDT).

In the PASAT (Gronwall, 1977), numbers are presented about every 3 s and the subject's task is to add them continuously. De Waard (1996), for example, used this secondary task to study the effects of car phones on driving workload. The *peripheral detection task* (PDT) was first introduced by Miura (1986). During driving, small visual stimuli are repetitively presented in the periphery of the subject's field of view (11 to 23 degrees of visual angle) for about 1 s. If the driver detects the stimulus he/she has to respond with a speeded key press. It is assumed that increased workload demands force the driver to focus their visual attention more narrowly to the line of sight and hence detection rates decrease and/or reaction times increase (Johansson et al., 2004). Martens and van Winsum (2000) used this technique in a simulator study and showed that PDT reflected changes in the workload demands when the driver encountered obstacles (e.g., a breaking car ahead). Olsson and Burns (2000) applied the PDT in a real driving situation and could demonstrate how the handling of a built-in radio or CD-player increased workload demands and decreased the PDT performance. Furthermore, Höger (2001) used a version of the task in which he presented small visual transients at various positions in the driver's field of view in order to probe their visual attention.

* Corresponding author. Department of Psychology, Ludwig-Maximilians-University, Leopoldstr. 13, 80802 Munich, Germany. Tel.: +49 89 2180 6229; fax: +49 89 2180 5211.
E-mail address: baldauf@psy.uni-muenchen.de (D. Baldauf).

0003-6870/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apergo.2009.01.004

Detection rates were increased if probes were presented at task-relevant locations (such as traffic signs or other cars) compared to probes that appeared at task–irrelevant positions. Also, probes that were presented at virtual larger distance were detected later than probes that were flashed in the foreground. Unfortunately, the PDT has some disadvantages. Jahn et al. (2003) pointed out that PDT performance is often affected by head and eye movements. Since the stimuli appear at different excentricities depending on these movements, which cause the line of sight to change, performance is hard to compare.

Hart (1975, 1978) suggested that the measurement of time perception could function as an effective indicator to measure the cognitive load of a primary task. In several studies, *prospective timing* tasks (i.e., subjects are aware of the timing task) have been used in experimental settings, including flight simulations to assess pilots' mental workload (Hart, 1978; Casali and Wierwille, 1983; Wierwille and Connor, 1983; Zakay and Shub, 1998). Most of these studies show that prospective timing is a non-intrusive and highly sensitive secondary task reflecting current mental workload. Hart (1978) and Zakay and Shub (1998) found that pilots consistently underestimated time intervals when there was a greater task load; increasing task difficulty caused the length of produced intervals to increase.

Conceptual models, such as the attentional-gate model of prospective duration judgement (Thomas and Weaver, 1975; Zakay, 1989; Zakay and Block, 1996), propose that temporal judgements are based on at least three processing stages. The first component consists of an oscillatory pacemaker emitting pulses at a constant rate. These pulses are gated into an accumulator when a switch is closed, i.e., while an interval is being processed. The content of the accumulator provides the raw material for measuring time because the number of pulses accumulated during the time interval corresponds to perceived duration. Therefore, the time estimates depend on the number of pulses (time units) that were accumulated during the test interval. The output of the accumulator is transiently stored in a working-memory system for comparison with the content of reference memory, which contains a memory representation of pulses accumulated in past situations. Finally, at a decision level, a mechanism compares the current duration values with those in working or reference memory to arrive at an adequate temporal response.

The assumed switch, which regulates whether continuously emitted time pulses are accumulated or not, is thought to be sensitive to central processing demands and cognitive workload because both the temporal and the non-temporal (primary) task compete for attentional resources (Brown, 1996). The result of an increased allocation of attention to a non-temporal task is that fewer time pulses accumulate in the counter as compared to a situation in which a subject directs primary attention to time. In highly demanding situations there are less mental resources available for the described timer mechanism and therefore less time pulses are stored in the accumulator. If there are fewer pulses in the *accumulator* under demanding conditions, a time-estimation task will result in an *underestimation* of objective time (Zakay, 1993; Block and Zakay, 1997; Wittmann, 1999) because less pulses have been accumulated compared to the standard in the reference memory. In contrast, increased attention to a non-temporal task leads to a relative *overproduction* of an interval because it takes longer for a sufficient number of pulses to be collected. This attentional model has been empirically supported in many studies (McClain, 1983; Fortin and Rousseau, 1984; Brown and West, 1996; Zakay and Shub, 1998).

The aim of the present study was to test whether a time-production task could be used as a valid indicator of mental workload in a driving simulator. In this study, prospective timing was embedded in a dual-task condition. The primary task was to drive a car in a virtual environment. As a secondary task, participants had to actively produce a time interval with certain duration. According to the discussed timing models, produced time intervals should be prolonged under driving conditions with higher workload.

To evaluate the suitability of time production as an indicator of mental workload, we simultaneously recorded a physiological parameter, electrodermal activity (EDA), which has served as an indicator for mental workload in several studies (e.g., Wilson and O'Donnel, 1988; Wilson, 2001). A third parameter, subjective ratings via the *subjective workload assessment technique* (SWAT; Reid and Colle, 1988), was recorded. The comparison of a physiological parameter, subjective ratings, and a behavioral secondary task should reflect the conceptual complexity of cognitive workload demands (O'Donnel and Eggemeier, 1986). By exposing the participants to a continuous three-part task with different complexity, we manipulated the demands on drivers' mental workload. During a 'straight road' section, the driver's mental workload should be relatively low because driving on the straight road with almost no curves and without any obstacles should be comparably effortless. In contrast, another section with curves and 'oncoming traffic' should require more attention resources. Not drifting over the markings on the right and avoiding collisions with oncoming traffic demand higher levels of concentration in this part of the simulation. A qualitatively different section is the third one, in which participants have to drive through a virtual city. This requires a relatively high mental workload because a high degree of visual attention and quick reactions are necessary to avoid accidents with suddenly appearing obstacles.

2. Method

2.1. Apparatus

For the driving simulation we used a static-base driving simulator from Dr. Foerst AG, Germany, without vestibular feedback. With this simulator we were able to study driving behavior in conditions that closely approximate those in real life without the risks of actually driving on the road. In contrast to real driving, the events occurring during the test drives can be independently configurated. A motorless compact car (Smart) with automatic gear shift was assembled in the simulator, and several car functions were recorded via sensor systems. These functions included data from the ignition, as well as brake and throttle control, which were transfered to the simulator's software to provide and control the car's movements in virtual reality depending on the driver's actions.

The simulated visual environment was generated via six image generators with a refreshing rate of 40 Hz which were interconnected and synchronized via a computer server. Three of the generated images were projected by a video projector above the car roof onto three big silver screens (each 2.05×2.05 m, 3.30 m distance from the car) in front and on the side of the vehicle to provide the driver with straight-ahead and side views with 120 degrees of visual angle in excentricity. The other three images were displayed on big monitors (1.50×1.00 m) standing behind the car and served as the rear view in the side mirrors. All image generators were synchronized in such a way that objects appearing in front of the car smoothly slid along the side into the rear-view mirror. Also sounds from the motor and environment were simulated, and the speedometer display showed the speed of the car in its simulated environment. During simulation a variety of driving data, such as frontal and lateral acceleration, were recorded with a refreshing rate of 20 Hz. The simulator was also equiped with an additional channel through which trigger signals were transferred to a separate computer that controlled the recordings of physiologic parameters.

2.2. Driving simulation

For our study we designed a round course (7.2 km) that could be divided into three main sections – ‘city’, ‘straight road’, and ‘oncoming traffic’ – which were connected by neutral intercepts (slight curves and straight roads) and created different demands on cognitive workload. In this set-up, we added a variety of different traffic events like oncoming cars, pedestrians, children playing, etc. A drive was defined as performing the whole round course including all three sections. The term ‘trial’ refers to performing a single time-processing task during one single section.

In the ‘city’ section, the driver passed through a town in which there were many events demanding caution and appropriate reactions by the driver. In particular, there were five special events on a road length of 1200 m. A person in a parked car opened his door shortly before the subject’s car passed. At a traffic light one had to give way to a car coming from the left-hand side and wait about one second until the lights turned green. Furthermore, there was a pedestrian who crossed the street 30 m in front of the car, another car approached from the right-hand side with right-of-way at the second crossing (without a traffic light), and, finally, a child with a ball jumped out from behind a parked car 40 m ahead. All these events were rather homogenously spread throughout the ‘city’ section. In the ‘straight road’ section there was no special event, and the driver simply had to keep the car straight on the road with almost no curves. Here, frequent landmarks, like trees, were visible. In the third section with ‘oncoming traffic’, the driver had to maneuver along a winding road. A stream of oncoming cars appeared at the first curve of the section and continued passing until its very end. The speeds of these cars ranged from 60 to 80 km/h. The number of oncoming vehicles was rather high, and the distance between the vehicles varied.

The same scenarios were used for all subjects. The starting point of each drive was set randomly at one of the three interconnecting intercepts, i.e., always 400 m before the beginning of one of the above specified experimental sections. The subjects were instructed to drive in the middle of the right lane and to maintain a continuous speed of 70 km/h except in the city section, in which they had to limit their speed to a maximum of 50 km/h (the usual speed limit in German cities). All three sections were adjusted in length so that the time required to perform them was approximately the same (about one minute). At the end of each section subjects were requested to stop the car for the assessment of subjective workload (see workload-related parameters). An initial test drive at the beginning of the experimental session served as baseline for driving performance.

2.3. Workload-related parameters

The various workloads induced by the different sections were expected to have effects on several parameters, which directly reflect mental workload.

2.3.1. Driving performance

As an indicator of driving performance we analyzed: (1) the *lateral deviation* from the center of the lane (to the left or to the right) and (2) *absolute the longitudinal acceleration* of the car (increasing or decreasing speed). Low values in these driving parameters indicate good steering control and stable and consistent driving (Chiang et al., 2004; Noy et al., 2004; Wittmann et al., 2006). These variables function as an indicator of driving task difficulty.

2.3.2. SWAT

The subjective rating of cognitive workload was assessed with a German translation of the Subjective Workload Assessment Test (SWAT, Reid et al., 1981; Reid and Nygren, 1988) directly after finishing each of the three sections. In the SWAT, mental effort required in a situation is rated on three subscales: temporal density of events (time load), required concentration (mental-effort load), and subjective feeling of emotional stress and anxiousness (stress load). Each subscale is rated on three discrete levels (1–3, 3 being the highest value).

2.3.3. EDA

The electrodermal activity (EDA) was recorded to obtain a physiological parameter which is known to be a reliable measurement of cognitive workload (Wilson, 2001). In the analysis, the tonic component of EDA was used that reflected baseline shifts in absolute skin potential level. The subjects’ electrodermal activity was registered by two electrodes on the palm of the left hand (dermatom C6 and C8). The electrodes were placed so that they didn’t disturb the driver and weren’t affected by the steering movements. The data stream was synchronously recorded under dual-task condition (test drive) by a program on an additional computer system (Neuroscreen by Jaeger–Toennies) that received trigger signals from the simulation software. To quantify the tonic changes, we computed the change in the level of electrodermal signals relative to the preceeding neutral intercept. All three critical sections were interconnected by such a neutral intercept, which was characterized by a smoothly winding track and no special events.

2.3.4. Time perception

Participants were asked to produce a time interval while they were driving in a certain section. At the beginning of each section, an acoustic start signal was given, and the subject had to press a button mounted below his/her right index finger on the steering wheel as soon as s/he thought that the fixed interval of 17 s had elapsed. This is referred to as a time-production task. Zakay and Shub (1998) call this procedure current duration production (CDP). There is evidence in the literature – mainly basic-research studies or flight-simulation studies – that this kind of secondary task does not interfere with the performance of the primary task (Zakay and Shub, 1998). In many previous flight-simulation studies, CDP was found to be sensitive to mental workload demands (Casali and Wierwille, 1983; Wierwille and Connor, 1983; Casali and Wierwille, 1984; Bartolussi et al., 1986).

2.4. Procedure

Before the actual experiment began, subjects had the opportunity to become familiar with driving in the simulator and the time-production task. They first performed a training drive on a round course identical to the one in the experiment. This drive was used to obtain baseline driving parameters of lateral deviations and longitudinal accelerations. Following this, they received training in the time-production task while driving straight on a simulated highway without any traffic. While driving at a constant speed, they were asked to produce various time intervals (10 s, 25 s, and 35 s), which were different from the one used in the following experimental drive. In this training drive, subjects got feedback about the accuracy of their time production. We used this training setting to ensure that participants were at a similar performance level and to be able to reduce baseline variations between good and bad time estimators. Thus, we tried to enhance the accuity and resolution of

Please cite this article in press as: Baldauf, D., et al., Time perception as a workload measure in simulated car driving, Applied Ergonomics (2009), doi:10.1016/j.apergo.2009.01.004

the time-processing instrument as an objective indicator of mental workload. The training session took about three minutes.

For the experimental drive the subjects were instructed to concentrate on the traffic and to avoid any collisions or accidents. Electrodermal activity was recorded during the entire drive. Additionally, the time-production task had to be performed in each of the three sections, namely the city, straight road, and oncoming-traffic section (the order of sections was randomized across subjects). For this secondary task, an acoustic signal was presented shortly after the beginning of each critical section. Starting with that signal, subjects were instructed to actively produce a fixed interval of 17 s without counting. As soon as the driver thought the requested amount of time had elapsed, s/he had to terminate the interval by pressing a button mounted on the steering wheel. To prevent the subjects from silently counting the seconds, a distracting question was asked during each section, like, "How fast are you driving right now?" or "Which value is on the distance indicator?" To answer such questions, subjects had to read an instrument in the simulator and report the value verbally. In the middle of each section the subjects had to answer just one such easy question. A single question is distracting enough to prevent the driver from using a counting strategy and it does not substantially interfere with the primary task. The subjects had no difficulties in answering these questions. Most of them reported afterwards that it was impossible to count under these conditions. After each section the subjects were then asked to rate the cognitive demand of the last section on the three SWAT subscales.

2.5. Participants

Sixteen (10 female and 6 male) adults were paid for their participation in the experiment. Their ages ranged from 20 to 54 years (mean 34 years). They all had normal or corrected-to-normal vision and were right handed. All participants had valid drivers' licences and at least two years of driving experience (mean = 15.4, SE = 2.3). The participants were requested to remove their watches for the duration of the experiment.

3. Results

3.1. Driving performance (primary task)

All 16 subjects performed well in the primary task without any major errors, like collisions or departures from the road. The black bars in Fig. 1a and b provide an overview of the means of the two measured driving parameters, the lateral deviation from the road, and the longitudinal acceleration in the three drive sections. Fig. 1a indicates that it was much harder for the participants to keep the car in the middle of the lane under the more challenging driving conditions 'city' and 'oncoming traffic' compared to the more stable driving in the condition 'straight road'. In the test drives (without the secondary time-production task) similar values were observed, as indicated by the striped bars in Fig. 1. For the lateral deviation a two-way ANOVA showed a significant main effect of the factor 'drive section' ($F[2, 30] = 7.01$, $p < 0.003$), but no effect of the factor 'task' with the levels 'dual task performance' versus 'driving only' ($F[1, 15] = 1.34$, $p > 0.26$) and no interaction ($F[2, 30] = 1.74$, $p > 0.19$). Post-hoc *t*-tests for the drives under dual-task conditions revealed significant differences in the lateral deviation of the car between the 'city' and the 'straight road' condition ($t[15] = -5.10$, $p < 0.001$), as well as between 'oncoming traffic' and 'straight road' ($t[15] = -7.9$, $p < 0.001$, all *p*-values are Bonferroni adjusted). Performance in the two highly demanding sections 'city' and 'oncoming traffic' did not significantly differ ($t[15] = -1.23$, $p > 0.23$).

A similar pattern of results was observed for the longitudinal acceleration of the car (black bars in Fig. 1 b). An ANOVA showed a significant main effect only for the factor 'drive section' ($F[2, 30] = 14.2$, $p < 0.001$), but not for the factor 'task' (dual task performance versus driving only, $F[1, 15] = 1.13$, $p > 0.30$) and no interaction between the two factors ($F[2, 30] = .98$, $p > 0.38$). The longitudinal acceleration of the car was on average significantly greater in the 'city' section than under the 'straight road' condition ($t[15] = -5.18$, $p < 0.003$). No statistical difference could be observed between the conditions 'city' versus 'oncoming traffic' or between 'straight road' versus 'oncoming traffic' ($t[15] = 1.41$, $p > 0.17$ and $t[15] = -1.65$, $p > 0.11$, respectively). In the two more demanding sections ('city' and 'oncoming traffic'), most participants had to readjust the speed of the car more often by moderate acceleration and deceleration. In contrast, participants were better able to maintain a constant speed under the straight road condition.

3.2. Time production

The length of the produced time interval varied according to the cognitive load in each drive section. Fig. 2 shows the mean length of produced intervals for each part of the round course. As can be

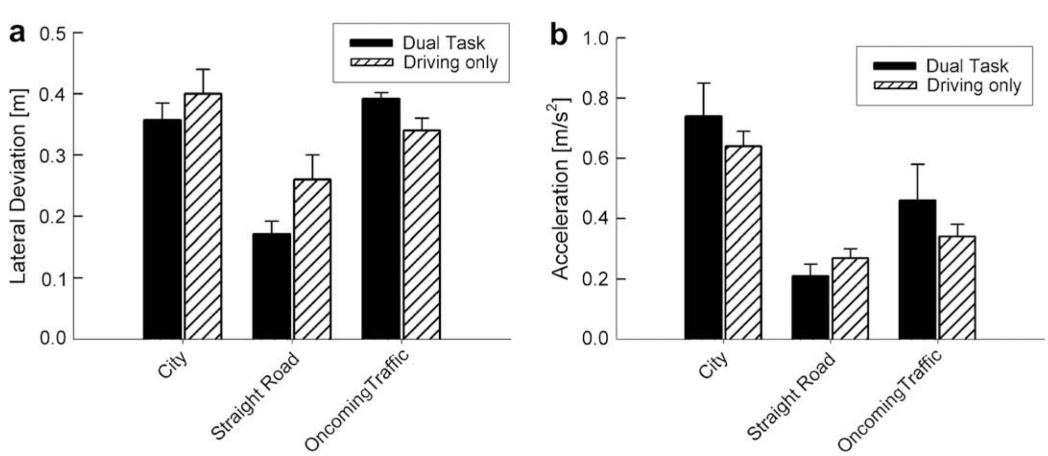

Fig. 1. Performance in the primary driving task. (a) The black bars represent the mean lateral deviation from the centre of the lane as a function of the experimental driving conditions. The striped bars show the same parameter during the test drive without the simultaneous time-production task. (b) The absolute longitudinal acceleration of the car as a function of the driving conditions (means and standard errors).

Please cite this article in press as: Baldauf, D., et al., Time perception as a workload measure in simulated car driving, Applied Ergonomics (2009), doi:10.1016/j.apergo.2009.01.004

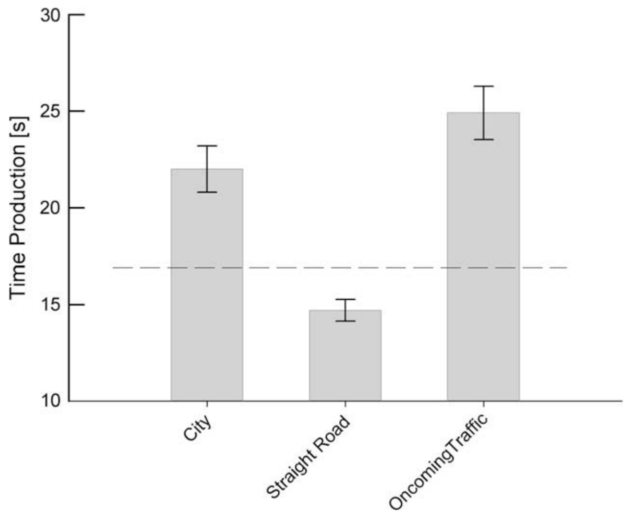

Fig. 2. Productions of 17-second intervals in the secondary task as a function of the three driving conditions. The horizontal line marks the real 17s-interval that should be produced (means and standard errors).

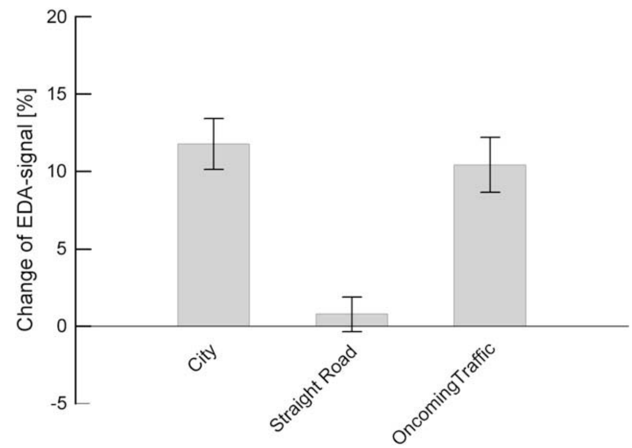

Fig. 3. Percent change in EDA signal relative to the preceding neutral intercept (means and standard errors).

seen, the more demanding 'city' section coincided with an overestimation of time (a mean of 21 s for the 17-second interval). In contrast, the monotonous stretch in the 'straight road' section led to a slight underestimation of absolute time (a mean of 15 instead of 17 s). In the section with 'oncoming traffic', participants again overestimated the time interval (on average 24 s for the 17-second interval).

A one-way ANOVA shows a significant main effect of the factor 'drive section', $F(2, 30) = 35.48$, $p < 0.01$. Pairwise *t*-tests show that the produced time of both highly demanding sections ('city' and 'oncoming traffic') was significantly longer than in the less demanding 'straight road' section, $t(15) = -4.43$, $p < 0.001$ and $t(15) = -6.23$, $p < 0.001$, respectively. There was only a marginally significant difference between the time productions in the two highly demanding sections, $t(15) = -1.80$, $p > 0.08$.

3.3. *Electrodermal activity*

Electrodermal activity as a physiological parameter also varied significantly with changing driving demands. Fifteen out of sixteen subjects showed an increase in EDA level in the 'city' section of more than 2 microvolt, as did 15 out of 16 in the 'oncoming traffic' section. Six out of 16 participants showed only slight increases (less than 2 microvolt) during the less demanding section ('straight road'), and five subjects' electrodermal activity even decreased during this part of the course relative to the preceding neutral intercept.

An ANOVA shows a significant main effect of factor 'type of section' on the change in electrodermal baseline activity, $F(2, 30) = 21.93$, $p < 0.001$. Fig. 3 shows the mean shift in baseline activity (in percent) in relation to driving sections. Pairwise *t*-tests revealed significant differences between 'city' and 'straight road' ($t[15] = -4.99$, $p < 0.001$), as well as between 'oncoming traffic' and 'straight road' ($t[15] = -4.22$, $p < 0.001$), but not between 'city' and 'oncoming traffic', $t(15) = -0.78$, $p > 0.44$ (all *p*-values are Bonferroni adjusted).

3.4. *Subjective ratings*

An ANOVA delivers a significant main effect of the factor 'section' on subjective ratings in the SWAT scale sum, as well as in all three subscales, $F(2, 30) = 19.374$, $p < 0.001$. The highest mean SWAT score occurred in the 'city' section (2.3 on average over all subscales, $SE = 0.13$). The section with 'oncoming traffic' also obtained high scores on all three subscales (2.1 on average, $SE = 0.16$) (see Fig. 4). However, as expected, the participants rated the 'straight road' condition to be by far less demanding (1.25 on average over all subscales, $SE = 0.08$).

Pairwise *t*-tests showed significant differences between the total scores of sections 'city' and 'straight road' ($t[15] = -5.65$, $p < 0.001$) and between 'oncoming traffic' and 'straight road', $t(15) = -4.49$, $p < 0.001$. No significant effect was noticed between 'city' and 'oncoming traffic', $t(15) = -1.15$, $p > 0.25$. The same effects could also be observed on all three subscale levels.

4. Discussion

The aim of the present study was to see whether a time-production task could be applied as a mental workload measure in a simulated driving environment. The study was embedded in a framework with a high level of ecological validity since it implemented driving behavior in a real car in a realistic yet

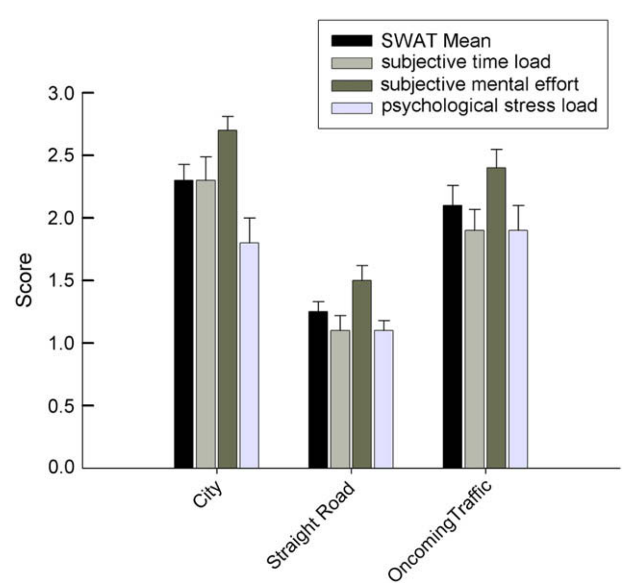

Fig. 4. SWAT mean score and the three subscales as a function of the driving conditions (means and standard errors).

Please cite this article in press as: Baldauf, D., et al., Time perception as a workload measure in simulated car driving, Applied Ergonomics (2009), doi:10.1016/j.apergo.2009.01.004

simulated driving environment. In the specific experimental settings that were used, the cognitive demands were manipulated within one driving session. The analysis of driving parameters showed that the three experimental conditions actually differed in experienced difficulty. In the conditions 'city' and 'oncoming traffic' it was much harder to keep the right lane and to drive at a constant speed. As the workload demands increased, the difficulty levels in our driving simulation were reflected in subjective ratings of workload, in the change in electrodermal activity, and in time production. Most importantly for the purpose of this study, the secondary time-production task did not significantly affect driving performance. This was revealed by the comparison between the driving performance under dual-task versus under single-task (test drive) condition (see black versus striped bars in Fig. 1).

The time-production task was found to be sensitive to changes in workload level. This is evident in the significant differences between the duration of produced time intervals in different sections of the course. The more demanding a situation was the more the participants tended to overestimate the time interval. These results support the attentional-gate model of prospective duration judgement (Thomas and Weaver, 1975; Zakay, 1989; Zakay and Block, 1996). The model assumes that under higher workload conditions a subject is distracted more from attending to time. The production of a certain time interval then takes longer because hypothesized time counts (or pulses) that represent that duration take longer to accumulate.

Both physiological data and subjective ratings indicate that the 'city' and 'oncoming traffic' sections were more demanding than the 'straight road' section. All three SWAT subscales delivered unambiguous data. Together with the recorded EDA stream, they can serve as a standard with which to validate time perception as a workload indicator.

The EDA recordings were transformed to a value indicating percent of change in signal to make them comparable among all subjects independent of individual baseline differences. In most subjects, electrodermal activity increased strongly with the higher demands of a section. EDA remained unchanged or even decreased relative to neutral intercepts when cognitive demands were at the lowest level during the 'straight road' section. Since the baseline EDA signal was recorded under the dual-task condition of the test drives the baseline might have been affected by an increased activation level. The possibly higher level of activation in the dual-task condition compared to the 'driving only' condition may have induced a ceiling effect. As a consequence, subtle differences between the two highly demanding conditions ('city' and 'oncoming traffic') may have been masked.

Time production parameters were comparably sensitive to the experimental manipulation of the driving situation. This may indicate that all of them measure similar changes in cognitive demands and provide converging evidence for the suitability of time production as a measurement of mental workload. The findings obtained in the present experiment are consistent with findings from flight simulations in former studies, in which prospective time-estimation tasks were employed (Casali and Wierwille, 1983; Wierwille and Connor, 1983; Casali and Wierwille, 1984; Bartolussi et al., 1986).

In our study we used a fixed interval of 17 s for the time-production task. Previous studies have successfully used similar time intervals for secondary time-production tasks as workload measures in laboratory as well as in simulator studies (e.g., Zakay and Shub, 1998). Time intervals in the range of 15–30 s are long enough to easily prevent participants from counting strategies. Longer intervals in the range of minutes, however, may increase the variability in time production. The length of the interval is crucial and strongly depends on the characteristics of the primary task. In order to use time production procedures as workload assessment in a given primary task the interval should be shorter than the driving section of interest. Further research could determine the optimal time interval length that is used to assess workload demands in various primary task conditions. For example, it could be of special interest whether a reduction of the test interval that has to be produced in the secondary timing task increases the temporal resolution of the workload assessment. This could make the time-production task an even more powerful technique and enable the differentiation between more similar driving situations such as the two demanding conditions in the present study.

The proposed time-production task has several specific advantages as workload measure in complex situations such as driving a car. In comparison to physiological measures, like the EDA signal recorded on the participant's hand, it is much easier to collect by registering single button presses. Although the EDA signal is known to highly correlate with task demands and the activation level of participants, it is comparably complex to record and to analyse those data online. Physiological measures are also affected by movement artifacts, which are hard to control for in natural environments like cockpits. Moreover, physiological recordings are prone to temperature artifacts and also require a reliable baseline score for interpretation.

In comparison to secondary task measures like the *peripheral detection task* (PDT) it is advantageous that the time production specifically interferes with the *cognitive* workload, whereas PDT can not differentiate between visual distraction and cognitive workload (see Johansson et al., 2004). Especially in situations like driving, subjects have to actively navigate and therefore frequently move their head and eyes. Furthermore, complex traffic situations require not only head and eye movements but also covert shifts of attentional focus, for example, in order to track multiple objects. In real driving situations, in which the PDT stimuli are projected to the windscreen, detection rates in the PDT – measurement are crucially affected by the location of eye fixations and where the subject may be covertly attending to while the visual stimulus is presented. The visual detection of probe stimuli is further affected by lighting conditions, which are hard to control for under natural driving conditions. The more implicit method of measuring mental workload through the perception of time could be helpful when trying to avoid subjects' biases in explicit answers (e.g., in inventories like the SWAT) due to expectations.

To sum up, we found the concurrent time-production task to be a valid indicator of mental workload in simulated car driving. The differences in time production were complemented by similar differences in parameters, such as electrodermal activity and subjective reports of workload. Numerous studies have shown that the experience of time is an indicator of cognitive functioning. Specifically, the assessment of time perception, as we have demonstrated here, could become a tool in applied ergonomics.

References

Bao, Y., Kiss, M., Wittmann, M., 2002. Effects of age and memory grouping on simulated car driving. In: Proceedings of the Human Factors and Ergonomics Society 46th Annual Meeting, pp. 1853–1857.

Bartolussi, M.R., Kantowitz, B.H., Hart, S.G., 1986. Measuring pilot workload in a motion base trainer. Appl. Ergon. 17, 278–283.

Block, R.A., Zakay, D., 1997. Prospective and retrospective duration judgements: a meta-analytic review. Psychon. Bull. Rev. 4, 184–197.

Brown, S.W., 1996. Attentional resources in timing: interference effects in concurrent temporal and non-temporal working memory tasks. Percept. Psychophys. 59, 1118–1140.

Brown, S.W., West, A.N., 1996. Multiple timing and the allocation of attention. Acta Psychol. 75, 103–121.

Casali, J.G., Wierwille, W.W., 1983. A comparison of rating scale, secondary-task, physiological, and primary-task workload estimation techniques in simulated flight task emphasizing communications load. Hum. Factors 25, 623–641.

Casali, J.G., Wierwille, W.W., 1984. A comparison of rating scale, secondary task, physiological, and primary-task workload estimation techniques addressing sensitivity and intrusion issues. Ergonomics 27, 1033–1050.

Chiang, D.P., Brooks, A.M., Weir, D.H., 2004. On the highway measures of driver glance behavior with an example automobile navigation system. Appl. Ergon. 35, 215–223.

O'Donnel, R.D., Eggemeier, F.T., 1986. Workload assessment methodology. In: Boff, K.R., Kaufman, L., Thomas, J.P. (Eds.), Handbook of Perception and Human Performance. Cognitive Processes and Performance, vol. 2. Wiley, New York, pp. 42.1–42.49.

De Waard, D., 1996. The Measurement of Drivers' Mental Workload. PhD thesis, University of Groningen, Traffic Research Centre, Haren.

Fortin, C., Rousseau, R., 1984. Time estimation as an index of processing demand in memory. Percept. Psychophys. 42, 377–382.

Gopher, D., Donchin, E., 1986. Workload – an examination of the concept. In: Boff, K.R., Kaufman, L., Thomas, J.P. (Eds.), Handbook of Perception and Human Performance. Cognitive Processes and Performance, vol. 2. Wiley, New York, pp. 41.1–41.49.

Gronwall, D., 1977. Paced auditory serial-addition task: a measure of recovery from concussion. Percept. Mot. Skills 44, 363–373.

Hancock, P.A., Mechkati, N., 1988. Human Mental Workload. North-Holland, Amsterdam.

Hart, S.G., 1975. Time estimation as a secondary task to measure workload. In: Proceedings of the 11th Annual Conference on Manual Control. US Government Printing Office, Washington DC, pp. 64–77.

Hart, S.G., 1978. Time estimation as a secondary task to measure workload: summary of research. In: Proceedings of the 14th Annual Conference on Manual Control (NASA CD-2060). NASA, Washington, DC, pp. 82–90.

Hart, S.G., Wickens, C.D., 1990. Workload assessment and prediction. In: Booher, H.R. (Ed.), An Approach to Systems Integration. Van Nostrand Reinhold, New York, pp. 257–296.

Höger, R., 2001. The signal location task as a method quantifying the distribution of attention. In: Harris, D. (Ed.), Engineering Psychology and Cognitive Ergonomics, vol. 5. Ashgate, Aldershot, pp. 373–380.

Jahn, G., Oehme, A., Rösler, D., Krems, J., 2003. IHRA-ITS-Swedish–German Study on Method Development. Federal Highway Research Institute (BASt) Project Nr. FE 82.175/2000. Chemnitz University of Technology, Chemnitz, Germany.

Johansson, E., Engström, J., Cherri, C., Nodari, E., Toffetti, A., Schindhelm, R., Gelau, C., 2004. Review of existing techniques and metrics for IVIS and ADAS assessment. Adaptive Integrated Driver Vehicle Interface (AIDE) Product number: IST-1-507674-IP.

McClain, L., 1983. Interval estimation: effect of processing demands on prospective and retrospective reports. Percept. Psychophys. 34, 185–189.

Martens, M.H., van Winsum, W., 2000. Measuring Distraction: The Peripheral Detection Task. TNO Human Factors, Soesterberg, Netherlands.

Miura, T., 1986. Coping with situational demands: a study of eye movements and peripherial vision performance. In: Gale, A.G., Brown, I.D., Haselgrave, C.M., Smith, P., Taylor, S.H. (Eds.), Vision in Vehicles-II. Elsevier, Amsterdam.

Noy, Y.I., Lemoine, T.L., Klachan, C., Burns, P.C., 2004. Task interruptability and duration as measures of visual distraction. Appl. Ergon. 35, 207–213.

Olsson, S., Burns, P., 2000. Measuring distraction with a peripheral detection task. The National Highway Traffic Safety Administration Driver Distraction Internet Forum. On-line paper. Available at: http://www-nrd.nhtsa.dot.gov/departments/nrd-13/driver-distraction/Papers20006.htm#A6.

Reid, G.B., Shingledecker, C.L., Nygren, T.E., Eggemeier, F.T., 1981. Development of multidimensional subjective measures of workload. In: Proceedings of the 1981 IEEE International Conference on Cybernetics and Society, pp. 403–406.

Reid, G.B., Colle, H.A., 1988. Critical SWAT values for predicting operator overload. In: Proceedings of the Human Factors Society 32nd Annual Meeting, pp. 1414–1418.

Reid, G.B., Nygren, T.E., 1988. The subjective workload assessment technique: a scaling procedure for measuring mental workload. In: Hancock, P.A., Meshkati, N. (Eds.), Human Mental Workload. North-Holland, Amsterdam, pp. 185–218.

Thomas, E.A.C., Weaver, W.B., 1975. Cognitive processing and time perception. Percept. Psychophys. 17, 363–367.

Wickens, C.D., 1992. Engineering Psychology and Human Performance, second ed. Harper Collins, New York.

Wierwille, W.W., Connor, S.A., 1983. Evaluation of 20 workload measures using a psychomotor task in a moving base aircraft simulator. Hum. Factors 25, 1–16.

Wilson, G.F., 2001. An analysis of mental workload in pilots during flight using multiple psychophysiological measures. Int. J. Aviat. Psychol. 12, 3–18.

Wilson, G.F., O'Donnel, R.D., 1988. Measurement of operator workload with the neuropsychological workload test battery. In: Hancock, P., Meshkati, N. (Eds.), Human Mental Workload. North-Holland, Amsterdam, pp. 63–100.

Wittmann, M., 1999. Time perception and temporal processing levels of the brain. Chronobiol. Int. 16, 17–32.

Wittmann, M., Kiss, M., Gugg, P., Steffen, A., Fink, M., Pöppel, E., Kamiya, H., 2006. Effects of display location of a visual in-vehicle task on simulated driving. Appl. Ergon. 37, 187–199.

Zakay, D., 1989. Subjective time and attentional resource allocation: an integrated model of time estimation. In: Levin, I., Zakay, D. (Eds.), Time and Human Cognition: a Life Span Perspective. North-Holland, Amsterdam, pp. 365–397.

Zakay, D., 1993. Time estimation methods – do they influence prospective duration estimates? Perception 22, 91–101.

Zakay, D., Block, R.A., 1996. The role of attention in time estimation processes. In: Pastor, M., Artieda, J. (Eds.), Time, Internal Clocks and Movement. North-Holland, Amsterdam, pp. 143–164.

Zakay, D., Shub, J., 1998. Concurrent duration production as a workload measure. Ergonomics 41, 1115–1128.

Please cite this article in press as: Baldauf, D., et al., Time perception as a workload measure in simulated car driving, Applied Ergonomics (2009), doi:10.1016/j.apergo.2009.01.004