CHAPTER 4 – SHORT TERM CARDIOVASCULAR RESPONSES TO CHANGING TASK DEMANDS

CHAPTER 4 – SHORT TERM CARDIOVASCULAR RESPONSES TO CHANGING TASK DEMANDS

Abstract

When measuring operator states the predictive power of cardiovascular and respiratory measures in relation to mental workload has been questioned. One of the main questions is to what extent do cardiovascular measures actually reflect mental workload. This question arises because good measures of mental workload should be sensitive to changes in mental effort alone and not to other influences or at least the changes associated with mental workload should be easy to isolate. In the case of cardiovascular measures, the physiological change brought on by the baroreflex is a compensatory control effect that can potentially overshadow changes in physiology due to mental effort and therefore reduce the usefulness of cardiovascular measures. However, this does not need to be the case. Despite the effects caused by the baroreflex differences in heart rate, heart rate variability and other cardiovascular measures associated with task related effort can still be found using short-term response patterns. The short-segment analysis approach described in this paper is based on a time–frequency method in which the spectral power of the cardiovascular measures in specified spectral bands is computed from small time segments, i.e. 30 s. To demonstrate the effectiveness of this technique two studies which made use of a simulation of an ambulance dispatcher’s task are described, both with easy and difficult task conditions. A short-lasting increase in task demands was found to be reflected in short-lasting increases in heart rate and blood pressure in combination with corresponding decreases in heart rate variability and blood pressure variability. These effects were larger in easy task conditions than in hard conditions, likely due to a higher overall effort-level during the hard task conditions. However, the developed measures are still very sensitive to mental effort and if this brief segmentation approach is used cardiovascular measures show promise as good candidates for reflecting mental effort during the assessment of operator state

Introduction

Operator functional state assessment is increasingly becoming a major research topic especially in relation to adaptive automation or adaptive support functions (Hockey et al., 2003). In particular, research into the effects of mental effort on various cardiovascular and respiratory measures has been very valuable for the development of operator state assessment techniques (Wilson and Russell, 2003; Kramer, 1991; Veltman and Gaillard, 1996; Sirevaag et al., 1993; Prinzel et al., 2000; Grossman and Taylor, 2007; Wientjes, 1992; Mulder and Mulder, 1981; Stuiver and Mulder, 2009) and promising results have been obtained that show the predictive power of such measures in relation to mental effort in the field of adaptive automation (Haarmann et al., 2009; Fairclough and Venables, 2006). Ting et al. (2010), for example, presented a paper on real-time adaptive automation based on, amongst other measures, mid frequency heart rate variability data.
Specifically, there are two main research findings, which have been widely used when assessing operator state via cardiovascular measures. These are that heart rate measures tend to increase in relation to periods of additional effort during task performance and that heart rate variability (HRV) measures tend to decrease during these periods (Mulder, 1992). Interestingly however, there are also studies that show contradictory effects and inconsistencies in this relationship between mental effort and heart rate or heart rate variability, for instance, finding what appears to be an effort related decrease in heart rate instead of the expected increase (Veltman and Gaillard, 1996; Sirevaag et al., 1993; Mulder, 1992; Wilson, 1992; Jorna, 1992; Porges and Byrne, 1992).

These contradictory and inconsistent findings challenge the usefulness of psychophysiological measures to directly reflect the effects of mental workload. Additionally, as Mulder and Mulder (1981) have already stressed it is not task difficulty that is the relevant factor for the heart rate variability response but rather it is invested effort. An individual may choose not to react to increased task demands or adapt and use a different strategy to meet the demands. In either case this would result in no visible effect in any of the related psycho-physiological measures (Mulder et al., 2009; Meshkati et al., 1995).

Similarly, if performance is examined in (semi-)realistic working situations large fluctuations in both task load and invested effort may be expected without being directly visible in any indicators of task performance (Hockey, 1997).

Another question that arises from the contradictory findings mentioned earlier is to what extent is any one particular measure of mental effort sensitive and selective (Mulder, 1986; Vicente et al., 1987; Brookhuis and de Waard, 2010; Byrne and Parasuraman, 1996). While several studies have shown that relatively large heart rate and variability differences can be found between baseline rest periods and mental task performance (Veltman and Gaillard, 1996; Mulder et al., 1992), other studies have had difficulty showing any differences in these measures or even between the tasks themselves (Kramer, 1991; Wilson, 1992; De Rivecourt et al., 2008; Nickel and Nachreiner, 2003). One explanation for this could be related to the selectivity of the measures used. Mulder (1986) makes a distinction between effort related to cognitive processes (task related effort) and compensatory effort, which is related to compensation in reaction to non-optimal states. However, these two effects often occur at the same time and therefore can be confused with each other. So, if a measure is to be useful in detecting workload it must be able to only detect changes related to response to task demands and not to compensatory state control (De Waard, 1996).

Effort related to compensatory control processes is caused by ongoing physiological activation during task performance, which in many cases results in increased blood pressure. Although this is indeed an effect of mental effort it does not directly reflect the changes in the current task situation but is instead...
a result of sustained increased activation and the corresponding compensatory reaction of the body to restrict physiological effects over longer time periods, say from 15 min up to a few hours.

The typical initial response to workload is often called a defence response or a preparatory response for fight-or-flight and is associated with the activation of the sympathetic system. This sympathetic activation is in turn related to an increase in heart rate and blood pressure and a decrease in variability measures of heart rate and blood pressure (Mulder, 1992). Research on the initial cardiovascular responses examines the interactions between these responses and other relevant psychological factors such as mood and motivation (Obrist, 1981; Gendolla and Krüsken, 2001). However, when effort is required continuously for more than five to ten minutes the short-term blood pressure regulation system, also called the baroreflex system, starts to counteract the initial effects mentioned above (Mulder et al., 2009; Mulder, 1980; Jordan, 1990; Berntson et al., 1991). Van Roon et al. (2004) explain this phenomenon in terms of a stronger reacting short-term blood pressure control system, which is reflected in an increase of baroreflex sensitivity (Mulder et al., 2009). This results in, amongst other things, a decrease in heart rate and an increase in heart rate variability. Effectively this means that despite ongoing mental effort the increase in blood pressure is reduced and the effects in heart rate and HRV are opposite to the initial effects at the start of the mental workload increase. Therefore, this competition between the defence response and the effects of the baroreflex may be an important factor in explaining why contradictory effects of effort on heart rate and heart rate variability can be found during long lasting effortful task performance.

The present paper examines whether short-lasting increases in task demands can be effectively applied to study cardiovascular effects of mental effort. Given the short-time frame approach proposed, one issue that may arise is whether cardiovascular effects indeed occur and are strong and consistent enough on such a short time scale of seconds to minutes. In a semi-realistic experiment studying the relationships between mental effort, heart rate and heart rate variability over small intervals in a flight simulator De Rivecourt et al. (2008) found that analysis periods as short as 30 s gave reliable results if multiple short data-segments were averaged. In driving research this approach has been applied successfully to study the effects of increased effort investment when merging into traffic on a motorway (de Waard et al., 2008).

The short-segment approach presented in this paper is based on a time–frequency method in which the power in specified spectral bands is computed from small overlapping time segments as a function of time, by using 30 s moving windows. Short segments of this data can then be selected and related to specific events during an on-going cognitive task. The main hypothesis investigated in this paper is whether the presented short-segment approach is able to detect changes that are due to invested cognitive effort, and not compensatory control, as reflected in heart rate, heart rate variability and other cardiovascular measures during a long lasting (semi-)realistic task. To demonstrate that this is possible we used a simulation of an ambulance dispatcher’s task. The work of an operator in an
Chapter 4 – Short Term Cardiovascular Responses to Changing Task Demands

Ambulance dispatch centre is an engaging mental task that requires planning and the maintenance of good situation awareness (Blandford and Wong, 2004). During the task the operator has to manage and plan non-emergency patient transport to and from the hospital. Additionally, the operator receives emergency phone calls that they have to respond to by immediately sending the nearest non-occupied ambulance to the accident location. The emergency call can be considered to be an additional interrupting peak task load on top of the normal workload. Any cardiovascular response to the interrupting phone call can be considered as evidence for an effect of a short increase in cognitive mental effort.

Two studies, both of which use the ambulance dispatcher’s task, are reported which allow us to show whether the cardiovascular responses are consistent across studies via the short-segment method. It further allows us to demonstrate whether the visibility of short-lasting effort effects is dependent on the amount of effort already invested, as this factor differed between the two studies. It is expected that the short-lasting increase in mental effort due to dealing with the emergency call will be reflected in short-lasting increases in heart rate and blood pressure in combination with corresponding decreases in heart rate variability and blood pressure variability. It is also expected that these effects will be larger in easy task conditions than in hard conditions due to a higher overall effort-level in the reference time-segments that precede the emergency calls during the hard task conditions.

Methods

To evaluate the power of the short-term analysis method two studies were carried out in a simulated ambulance dispatcher’s centre. Because the methods of both studies are very similar most of the relevant details are described only once. However, where differences do exist they are clearly described.

Participants

The participants in both studies were students at the University of Groningen. They were given financial compensation for their participation and signed an informed consent form at the start of the first training session. Both studies were approved by the ethical committee of the Faculty of Behavioural and Social Sciences of the University of Groningen.

In the first study a total of 26 participants between the ages of 20 and 24 started the training and of this initial group 22 (12 female) finished the experimental sessions. The data of 21 of these participants were suitable for analysis. The loss of one participant in this case is because sometimes it is hard to get good blood pressure measurements, often because of slender or cold fingers. This causes excessive
arтеfacts in the blood pressure measurements and therefore such participants are removed from psychophysiological data analyses.

In the second study, a total of 22 participants aged between 19 and 27 completed the training and experimental sessions. After processing and checking for artefacts the data of 19 participants (12 female) were available for statistical analysis.

**Task**

A simulated ambulance dispatchers’ work environment was used in both studies. In this task trained participants had to handle emergency calls in order to send ambulances to accidents, handle non-emergency calls and schedule ambulances to pick-up patients at home and deliver them to the hospital and vice versa. Participants’ also had to make sure that there was enough coverage over the region, meaning that every location in the region was reachable by an ambulance within 15 min. The short increase of task load studied was the answering and responding to emergency calls that typically lasted 25 to 30 s.

The dispatchers’ task can be best described as a planning task with a continuous task load that was disturbed by incoming emergency calls. Within the three main tasks the operator had to perform the task with highest priority was activating rides as a response to emergency calls. During the emergency call participants had to select the best or most economic ambulance. This ambulance could have been one which was driving in the vicinity of the emergency or one from a post nearby. To be able to make this choice the operator needed a good overview of both the current locations and availability of the ambulances in the region.

The task with the next highest priority was making sure every location in the region could be reached by an ambulance within 15 min. To preserve this coverage the operator had to make sure he or she chooses the right ambulances for emergency rides and non-urgent transport, a task that required good insight into the locations and activities of the ambulances.

The third main task was scheduling the ambulances to carry out non-urgent transport rides to transport patients to and from the hospital for scheduled operations. Some of the planned rides were pre-sent in a list at the start of the scenarios and others were added to this list by the operator during the experiment via a non-emergency call. Planning and activating non-emergency rides required a lot of forward thinking and calculation. As such, the planning of non-urgent rides and maintenance of coverage in the region are considered to be the tasks that create the continuous task load that was occasionally interrupted or increased by the need to activate emergency rides.

In this study the dispatch centre was simulated on a computer with two 17 inch screens. On one screen the communication and planning interface was shown, which was operated by the participants with a
normal mouse and keyboard. On the other screen a map of the region was shown which displayed the ambulances currently in transit along with the locations of hospitals, ambulance stations, and emergency locations. A more detailed description of the simulation can be found in Mulder et al. (2009).

Training

In order to perform the task adequately participants first completed several extensive training sessions. The training focused on two main points. The first was the knowledge of the topography of the province used in the task. Participants were trained in remembering the location of most of the important locations in the region and the time it would take for ambulances to travel between these locations. However, if despite the training a participant could not remember where a city was located or how long it would take an ambulance to reach a destination during the experiment he or she could use an interactive map included in the task interface to access that information.

The second focus of the training was to make the participants familiar with the task, i.e. how to use the interface, find ambulances on the map, find cities and driving times between places, how to dispatch ambulances, plan non-emergency rides, and how to keep coverage. During this part of the training several specially developed scenarios were presented to the participants each of which trained them in different aspects of the task. Towards the end of the training period the scenarios became more complex and resembled the scenarios used in the experiment.

The participants in Study 1 received three two-hour training sessions. After the training their knowledge of the location of important places and the travel-times between those places was tested. Participants could only take part in the experimental session if they passed this test with a 60% accuracy rate. If they scored lower than 60% correct they were asked to practice some more and were re-tested prior to the experiment to make sure they had attained the 60% level. All participants reached the desired level before starting the experimental session.

The participants in Study 2 attended four two-hour training sessions. The level they had to reach was 80% accuracy for locations and 70% accuracy for travel-times. They were also tested after the training sessions and asked to study more and were re-tested prior to the experiment if their level was not high enough after the initial four training sessions. In Study 2 all participants also reached the desired level before starting the experimental session.

Experimental procedure

Study 1

For the first study the participants attended two three-hour experimental sessions during which they completed a total of 12 scenarios that lasted 15 min each. Four of these scenarios had an easy task level
where the participants only had to respond to nine telephone calls, activate four to five emergency rides, and maintain coverage of the region. During these easy scenarios the scheduled non-emergency rides were planned in such a way that they did not need to be activated during the scenarios at all and could effectively be ignored.

In the four difficult scenarios the participants also had to plan and activate four scheduled rides as well as respond to nine emergency calls and maintain coverage of the region. In the remaining four scenarios an advice system was tested. However, the results of these last four scenarios are not included in this paper. The different types of scenarios were presented in an alternating fashion with the presentation counterbalanced to minimize order effects.

The experiment itself was split into two sessions. The first session started with a five minute rest/baseline period after which six scenarios were presented. Measurements then continued for another 5 min after all of the scenarios were completed in order to provide a second baseline period. The second session with an own baseline period and the remaining six scenarios would either follow immediately after the first session or be carried out on a different day. Half of the subjects completed both sessions on one day and the other half on two separate days. If they received the second session on a different day this took place within a week of the first session and under similar conditions. The same day or different day split was made to check for fatigue effects on the task. No effects were found with respect to the split sessions and therefore the split is ignored for the rest of the analysis.

**Study 2**
In the second study participants completed a total of 16 scenarios (8 easy, 8 hard) that lasted 15 min each and were completed over two three-hour sessions. In the eight easy scenarios only four to five emergency rides had to be activated, coverage had to be kept and nine emergency and non-emergency calls answered. However, the non-emergency rides and incoming non-emergency calls were all designed in such a way that no ambulances needed to be activated for them at all during the scenario. As with the first study in the eight difficult scenarios four non-emergency rides had to be planned and activated within the time frame of the scenario.

**Cardiovascular measures**
In both studies cardiovascular measures were derived via an electrocardiogram (ECG) and by means of finger blood pressure measurements. For the ECG three Ag-AgCl electrodes were used. The common electrode was placed on the sternum and the other two electrodes on the right and left side of the chest between the two lowest ribs. Blood pressure was measured with a FIN.A.PRES device (Finometer®). Both ECG and blood pressure were recorded at sampling rate of 250 Hz and R-peaks were detected online from the ECG by using an ECG-trigger as part of a TMS Porti system (Twente Medical Systems International). Interbeat intervals (IBI’s) were then derived from the R-peak times and automatically corrected using CARSPAN (Mulder, 1992; Van Roon et al., 2004) followed by a visual
inspection to check if the corrected results were necessary, in which case they were accepted. Heart rate measures were computed as the number of beats per time unit (beats per minute).

Mean values of heart rate in beats/minute and systolic and diastolic blood pressure in mmHg were calculated in moving 30-s windows that were shifted in time by one-second steps. This means that the value at any set time point in the resulting profile data was calculated from a data-window starting from 15 s before to 15 s after that time point (Stuiver and Mulder, 2009). This method resulted in variables of heart rate and systolic and diastolic blood pressure as functions of time and that have new values every second.

Spectral values for heart rate and blood pressure, called heart rate variability (HRV) and blood pressure variability (BPV) respectively, were computed as natural log-transformed squared modulation index values from two spectral bands (see below). Heart rate variability and blood pressure variability as a function of time for each of these bands (profiles) were created by calculating spectral values from time segments of 30 s. For every data point in the variability measures the data from 15 s before that point in time and 15 s after were used to calculate a (spectral) value. The outcome of this process meant that the data for every time point overlapped by 29 s with the data for the next point. This method is very comparable to the method described above to calculate the mean values of heart rate, however, instead of calculating the average heart rate in the 30-s window a spectral band value was calculated for each of the two spectral bands and for both heart rate variability and systolic blood pressure variability. The measures that were obtained are called spectral profiles and represent the spectral power in the specified frequency band as a function of time (also called a time-frequency analysis). From these profiles an analysis of short-lasting time segments is chosen around an emergency telephone call. For baroreflex sensitivity a similar 30-s sliding window approach was applied, in this case however an estimation of baroreflex sensitivity was calculated in the 30-s periods.

Spectral analysis of all the cardiovascular data was performed with CARSPAN (Mulder, 1992; Van Roon et al., 2004). Spectral values were calculated for two different frequency regions. The first band ranged from 0.07 Hz to 0.14 Hz and is called the mid frequency band. The second, so-called (high) frequency band, ranged from 0.15 Hz to 0.40 Hz. This resulted in variability measures for heart rate and blood pressure in both frequency bands. Heart rate variability and blood pressure variability were calculated as values normalized to the mean, i.e. as a squared modulation index (Veldman et al., 1998). The variables were then logarithmically transformed to obtain normally distributed variables (Van Roon et al., 2004). An estimation of baroreflex sensitivity in ms/mmHg was computed by calculating the transfer gain (modulus) from changes in systolic blood pressure to changes in interbeat interval in the mid frequency range (Robbe et al., 1987).
Variables in the event related approach

In the introduction it was proposed that more insight into the effects of task demands can be obtained by looking at short periods of time (e.g. 30–60 s) instead of studying effects over a longer period of time (e.g. 15 min). Therefore in both of the studies described here short increases in task demands (20–30 s) were compared with preceding short periods of relatively low or average task demands. In this paper the period of increased task demands is referred to as the activation period and the period of low or average task demands that comes before the activation period is referred to as the reference period.

In both studies the activation period was tied to answering emergency telephone calls in the ambulance dispatch simulation. These calls, and the subsequent actions they require, typically last about 20–30 s during which the operator answers the phone, checks which ambulance is appropriate, and sends that ambulance to the accident location. As such, the value for the reference period was calculated as the average in the 20 s time segment just preceding the call. The corresponding activation period was calculated as the average from 10 to 30 s after they picked up to answer the call. The difference between the resulting reference and activation values is then taken as a measure for how strongly the person responded to the increased task demands. Also, by taking the averages of these values across a number of telephone calls the reliability of the measure is improved. Task load effects on effort investment can then be investigated by testing the interaction effect of workload condition and the reference period/activation period difference. To do so, for each participant and for each variable a value for the reference and activation period is calculated by averaging over all available events in a workload condition, which results in one value for the reference period and one for the activation period for both the easy and for the difficult conditions for each participant.

In the second experiment at the end of each scenario and before the next one started participants were asked to fill in the Rating Scale for Mental Effort (Zijlstra, 1993). This one item scale gives an impression of the mental effort participants experienced during task performance on a range from 0 to 150 with nine unevenly placed descriptive labels ranging from ‘absolutely no effort’ (a score of 2) to ‘extreme effort’ (a score of 112).

A General Linear Model Repeated Measures test (SPSS) was used to analyse the data. Repeated Measures ANOVAs were run with two factors (reference and activation) both with two levels (high and low workload). Meaning that in total the impact of workload, reference/activation, and the interaction of the two factors was analysed. Alpha was set to 5% for all of the tests.
Results

Study 1
The results for heart rate, systolic and diastolic blood pressure, and baroreflex sensitivity for study 1 are shown in Figure 7. The results for the spectral variability variables, i.e. heart rate variability and blood pressure variability are displayed in Figure 8. The results of the statistical analysis are summarized in Table 1. No effects of activation/reference or workload were found on average heart rate or for baroreflex sensitivity. However, systolic blood pressure ($F(1,20)=26.3$, $p < .001$) as well as diastolic blood pressure ($F(1,20)=9.50$, $p < .01$) increased during the activation period when compared to the reference period. In terms of the workload effects, diastolic blood pressure increased more in the low workload condition compared to the high workload condition, as expressed by the Ref/Act x Workload interaction effect ($F(1,20)=14.8$, $p < .01$). Additionally, lower values for heart rate variability in the mid-frequency band ($F(1,20)=8.14$, $p < .01$) were found in the high workload condition compared to low workload while a similar effect was observed for systolic blood pressure in the high frequency band ($F(1,20)=4.94$, $p < .05$).
The strongest main effects were found in the variability variables for the reference/activation difference along with their interaction with the workload condition. In general, heart rate variability showed a strong decrease during activation in the mid frequency band ($F(1,20)=29.0, p < .001$) as well as in the high frequency band ($F(1,20)=20.3, p < .001$). Blood pressure variability also decreased when the activation period started in both the mid frequency band ($F(1,20)=23.4, p < .001$) and in the high frequency band ($F(1,20)=26.6, p < .001$). All of the variability measures also showed an interaction effect between workload condition and the reference/activation difference. This means that heart rate variability showed a stronger decrease between the reference and activation period in the low workload condition compared to the high workload condition. This is apparent in both the mid frequency band ($F(1,20)=52.5, p < .001$) and in the high frequency band ($F(1,20)=25.6, p < .001$). Blood pressure variability showed the same interaction effect with a stronger decrease between the reference and activation period in the low workload condition for both the mid frequency band ($F(1,20)=17.0, p < .001$) and the high frequency band ($F(1,20)=26.2, p < .01$).
Figure 8: Reference/Activation differences in study 1 for HRV = Heart rate variability and SBPV = Systolic blood pressure variability in natural log-transformed squared modulation index values, in the mid and high frequency band. Reference and Activation correspond to the period before and during the task demands increase, respectively.
CHAPTER 4 – SHORT TERM CARDIOVASCULAR RESPONSES TO CHANGING TASK DEMANDS

Table 1: Results of study 1.

<table>
<thead>
<tr>
<th>Heart Rate</th>
<th>Systolic Blood Pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F(1,20)</td>
</tr>
<tr>
<td>Workload</td>
<td>0.40</td>
</tr>
<tr>
<td>Ref/Act</td>
<td>0.97</td>
</tr>
<tr>
<td>Workload *Ref/Act</td>
<td>0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diastolic Blood Pressure</th>
<th>Baroreflex Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F(1,20)</td>
</tr>
<tr>
<td>Workload</td>
<td>0.060</td>
</tr>
<tr>
<td>Ref/Act</td>
<td>9.50</td>
</tr>
<tr>
<td>Workload *Ref/Act</td>
<td>14.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Heart Rate Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Workload</td>
</tr>
<tr>
<td>Ref/Act</td>
</tr>
<tr>
<td>Workload *Ref/Act</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Systolic Blood Pressure Variability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Workload</td>
</tr>
<tr>
<td>Ref/Act</td>
</tr>
<tr>
<td>Workload *Ref/Act</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mid Frequency</th>
<th>High Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F(1,20)</td>
</tr>
<tr>
<td>Workload</td>
<td>4.03</td>
</tr>
<tr>
<td>Ref/Act</td>
<td>20.3</td>
</tr>
<tr>
<td>Workload *Ref/Act</td>
<td>25.6</td>
</tr>
</tbody>
</table>

Table 1: Results of study 1.

Study 2
The second study produced results comparable to the first, although, there are some differences. The results of heart rate, systolic blood pressure, diastolic blood pressure and baroreflex sensitivity can be found Figure 9 and the results of the variability measures are shown Figure 10. A summary of all of the statistical tests is given in Table 2. First of all, it is clear that heart rate was slightly higher in the high workload conditions than in the low workload conditions (F(1,20)=4.56, p < .05). Similarly, systolic blood pressure (F(1,20)=4.92, p < .05) and diastolic blood pressure (F(1,20)=6.02, p < .05) were also both slightly higher in the high workload condition. Main effects for the reference/activation difference were also present in heart rate and systolic blood pressure. This means that heart rate was lower during the activation period than during the reference period (F(1,20)=24.1, p < .001) whereas systolic blood pressure was higher (F(1,20)=9.04, p < .01). There was also a significant (F(1,20)=5.34, p < .01) interaction effect between reference/activation difference and workload in diastolic blood pressure, although the difference was less than 1 mmHg.
There was also a main effect of workload on blood pressure variability in the mid frequency band ($F(1,20)=5.19, p < .05$) indicating a lower HRV in the high workload condition. However, none of the other variability measures showed any main effects for workload. There were however, strong effects for the reference/activation difference in all variability measures, which were similar to those found in the first study. This means that heart rate variability decreased in the mid frequency band ($F(1,20)=35.9, p < .001$) and the high frequency band ($F(1,20)=6.41, p=.021$) during the activation period compared to the reference period. The same holds for blood pressure variability, which also decreased in the mid frequency band ($F(1,20)=24.0, p < .001$) and the high frequency band ($F(1,20)=12.0, p < .01$) during the activation period.
There were also interaction effects for the reference/activation condition and workload condition in the blood pressure variability measures. Blood pressure variability decreased between the reference and activation periods more in the mid frequency band in the low work-load condition than in the high workload condition (F(1,20)=9.45, p < .01) and BPV in the high frequency band showed a similar pattern (F(1,20)=5.29, p < .05). There were however, no interaction effects for the heart rate variability measures.

Results on subjective mental effort as measured by the RSME confirmed a difference in subjective effort ratings between the easy and difficult condition. Meaning that on average participants reported scores which were 16.6 points higher during the difficult condition (37.5 vs. 54.1, F(1,20)=49.50, p < .001).
Discussion and Conclusion

The main aim of this paper was to show that with a short-segment approach changes in invested effort could be detected during a long lasting (semi-)realistic task, as reflected in heart rate, blood pressure, heart rate variability and blood pressure variability. Such a technique would be useful as it would help with the development of operator state monitoring tools that are more sensitive to immediate changes in mental workload, i.e. effects of cognitive effort, and less sensitive to the effects of compensatory control. The results of the two studies presented in this paper seem to support the use of this short-segment approach. The results showed strong effects of workload manipulations on variability measures of heart rate and blood pressure as well as some effects on mean heart rate and systolic and diastolic blood pressure. The expected decrease of variability during a short workload increase, from a reference period to an activation period, was found in both studies for both the high and mid frequency bands of heart rate variability. The same effects were also found in both studies in both frequency bands in systolic blood pressure variability, which also shows the consistency of the data. Mean heart rate, blood pressure and baroreflex sensitivity measures, however, showed smaller or no

Table 2: Results of study 2.
effects, which could have been expected, as they are more related to the overall state of the cardiovascular system and less to immediate effects of changes in invested effort (Mulder et al., 2009).

A particularly interesting result was the interaction of task difficulty and the reference/activation difference measures. In the easier scenarios, where participants were not yet fully engaged, increasing workload for a short period of time showed a clear decrease in variability measures. Whereas, in the difficult scenarios the increase in workload from reference to activation did not provoke a further decrease of variability during the activation period, probably because variability levels were already low in the preceding reference period. This suggests that the participants were engaged already to a much higher degree during the difficult task condition and did not invest additional effort with the increase in workload during the emergency call.

The reference period is calculated from data that are in between or during normal task performance, making it different from a ‘normal’ baseline during rest. The values of the cardiovascular variables during these reference periods depend on what happens at that moment during task performance, which sub-tasks have to be performed, and what the current time pressure is. In situations where participants are already investing a lot of effort, little difference with the succeeding activation period may be expected. If on the other hand, the reference period is from a period where invested effort is low a difference with the activation period is likely since in this case the participants increase their effort in order to handle the emergency call. In this manner the difference between reference and activation may be informative in terms of a possible ceiling effect of workload. The magnitude of the response, or better a lack of response, may indicate that the participant is potentially reaching a workload redline (Brookhuis et al., 2003).

It should be noted that the above-mentioned interaction effect, i.e. a larger difference between the reference and activation period in the easy condition than in the difficult condition, is stronger in the first study compared to the second. In fact, for heart rate variability the inter-action effect is only present in the first study, while for blood pressure variability the effect is present in both but is stronger in the first study. Furthermore, in the first study the impact of additional workload was only present in the easy workload condition while it was present in both the easy and the difficult condition in the second study. This suggests the presence of a ceiling effect during the hard workload condition in the first study and the absence of one in the second.

The length of the scenarios, the number of actions, and the type of actions participants had to perform were quite similar in both studies, which gives little reason to expect the differences discussed above. Capacity theory implies that resources are scarce and that there is only a limited amount available, therefore suggesting the existence of a ceiling of effort. The amount of resources depends on things such as motivation and competence level (Hockey, 1997). The differences between the two current studies might therefore be explained by more training hours and a higher target score for
topographical knowledge in the second experiment. This may have resulted in better-trained participants that had a higher competence level at the moment the experiment started.

This leads to one of the limitations of this study, which is the absence of adequate performance data. To check the claims about whether the participants in the second study were better trained, performance data would have been very useful. Unfortunately the simulation of the dispatcher’s task does not provide accurate information about task performance that is comparable enough over the two studies and over the scenarios in order to be useful in the present context. It is therefore advisable for future studies to include more specific performance data in order to test the explanation given above for the difference in effects between the two studies.

It is worth noting that in the current analysis method an average was taken over multiple events (emergency calls) that were spread out over time. However, given the relatively small differences in task load between the easy and difficult conditions our results suggest that our measures are very sensitive when compared to previously described studies which used continuous measuring periods (e.g. Mulder, 1992; Kramer, 1991; Wilson, 1992; Nickel and Nachreiner, 2003). Further research is needed to confirm this claim.

In general it can be concluded that the short-segment measurement approach developed within this paper is sensitive to changes in workload even on a small time scale and despite simultaneously occurring compensatory effects. That cardiovascular spectral measures are usable on such small time scales confirms the findings of De Rivecourt et al. (2008). Furthermore, the current approach, after further research, could also be applied to create an index for mental workload, maybe in combination with other measures, in order to monitor and perhaps even assist operations. As such, the short-segment approach described in this paper seems to be a direction that future research for adaptive automation should consider.

Acknowledgements

Parts of this work were completed within the REFLECT project funded by the European Commission within the 7th Framework programme. The REFLECT project investigated ways of realizing pervasive-adaptive environment.