Human Factors Issues in Complex System Performance

Waard, D. de, G.R.J. Hockey, P. Nickel & K.A. Brookhuis (Eds.)

2007, Shaker Publishing, Maastricht, The Netherlands.

pp. 101 - 116.

Short-term heart rate measures as indices of momentary changes in invested mental effort

Ben Mulder, Heleanne Rusthoven, Marnel Kuperus, Michel de Rivecourt,
& Dick de Waard
University of Groningen
The Netherlands

Abstract

Laboratory research on invested mental effort in demanding mental tasks has given evidence for an increase in heart rate (HR) and decreased heart rate variability (HRV) as a function of task load. In particular the HRV effects are very consistent in short lasting tasks (e.g. 5 minutes) in which working memory is heavily involved. Using this concept in more practical workload situations evokes two types of problems: 1. Generally HRV increases as a function of time due to baroreflex related regulatory processes; 2. workload and mental effort change continuously as a function of time. This latter point is directly related to the reliability of the estimated HRV measures. According to published standards, in many cases time segments of steady workload would be too short to apply spectral techniques for HRV measures.

In this paper a simple time-frequency approach, called spectral profiles, for this type of problems is outlined that is suitable for many (semi-)realistic working conditions. This is illustrated in two studies using two different types of task, i.e. the SON IQ-test for children and a simulated flight for candidate pilots.

The results show that in both experiments HR and HRV effects based on short segment analysis can be related to task demands. It is concluded that the described short segment spectral profile approach is promising for use in (semi-)realistic working situations. The important pre-requisite is that on the basis of simple task analysis relevant segments have to be detected that do not overlap too much (in time or in effect) with respect of the HRV patterns.

Introduction

In several laboratory experiments it has been shown that heart rate variability (HRV) decreases with increasing task demands, while heart increases (Boucsein & Backs, 2000). In general a diminished HRV is interpreted as additional effort invested in the task to be performed (Mulder, 1986). On this basis HRV was proposed to be an index of invested mental effort. Many of these studies, however, can be characterised by short task duration (e.g. 5 minutes), while in most cases working memory load was manipulated specifically. Longer lasting tasks, in general, show a quite different

In D. de Waard, G.R.J. Hockey, P. Nickel, and K.A. Brookhuis (Eds.) (2007), *Human Factors Issues in Complex System Performance* (pp. 101 - 116). Maastricht, the Netherlands: Shaker Publishing.

pattern: after an initial HRV decrease at the start of a task, HRV increases gradually to levels even higher than preceding rest values (Mulder, van Roon, Veldman, Laumann & Burov, 2003) The question arises what the background is of these different patterns and how these can be explained.

The cardiovascular response pattern during short lasting tasks can be described as a defense reaction: compared to preceding rest situations heart rate and blood pressure are increased, while HRV, blood pressure variability (BPV) and baroreflex sensitivity are diminished. In many cases longer lasting tasks start with a similar pattern, but after five or ten minutes this pattern is followed by a distinctive decrease in heart rate in combination with an increase of baroreflex sensitivity and HRV. Mulder et al. (2003) interpret these changes as a response of the short term blood pressure control system (baroreflex) to the increased blood pressure. What happens with the blood pressure level depends on the task situation, but in many cases the initial increase in blood pressure is levelled off or the blood pressure level is even diminished.

It is clear that because of these different response patterns it is difficult to use HRV as a direct index of invested mental effort in long lasting tasks or during normal work in real life or even in simulated working tasks. In the literature it even led to conclusions that HRV is not always a valid index of invested effort during mental work (Boucsein & Backs, 2000, Nickel & Nachreiner, 2003). There are, however, more aspects that make it difficult to use HRV during normal daily work or in simulated workload studies. In normal daily mental loading work such as writing a report, planning, decision-making or car driving mental task load, and therewith required mental effort, is not constant but is varying from moment to moment. Short periods of peak load are interchanged with periods with less effort required. The same holds to a lesser extent for simulated real live situations, such as driving in a driving simulator, the ambulance dispatcher's task, or an air traffic control task. As a matter of fact four problems arise: 1. Is it possible to characterise complexity of specific task segments in terms of simple task analysis methods? 2. Can simple and complex task segments be traced in a simple manner on the basis of such a task analysis? 3. Are the mental loading segments long enough to evoke a distinct cardiovascular response pattern (i.e. increased HR and decreased HRV)? 4. Has an HRV-response on a preceding task event disappeared when a new event occurs, or is overlap between response patterns to be expected?

In the present paper an approach is outlined that tackles these problems. In particular, attention is given to the point of distinguishing short lasting task segments with varying task complexity and how such differences are reflected in HR and HRV patterns. As a matter of fact the question has to be answered whether it is possible to increase reliability of short segment spectral HRV-measures by taking together data of comparable task segments using averaging techniques. After the description of the spectral procedure that is used the results of two studies are described in the remainder of this paper, one related to HRV patterns during an intelligence test with children, the other related to workload in a flight simulator.

Spectral profiles analysis

The basis of the approach is a simple type of time-frequency analysis of HRV. In general in signal processing, time-frequency analysis is a method that delivers spectral distribution measures as a function of time. For instance, the time course of spectral power in a chosen frequency band can be computed or coloured 3-dimensional graphs can be plotted of the spectral power distribution, with time on one axis, frequency on another while the spectral power in a specific time window and specific frequency range is indicated by a colour scale. Sophisticated methods, such as wavelet-analysis procedures, exist in which the time window is adapted to the frequency range at hand: the lower the frequency range, the longer the time window. On this basis Van Steenis (2002) has developed a nice procedure that is suitable for HRV-analysis. For years we have used a simplified method in which the spectral power in a chosen frequency band as a function of time (spectral profile) can be computed (G. Mulder, 1980, L. Mulder, 1992). In this procedure the length of the time window can be adapted to the required frequency range, in such a way that the lower the frequency range of the band, the larger the chosen time window. Theoretically, the time window has to have at least the length of the inverse of the lowest frequency involved. Practically and for reasons of reliability of the computed spectral values, longer lasting time periods are chosen in most situations. For instance, when the power in the frequency band ranging from 0.07 – 0.14 Hz is selected (mid frequency band of HRV) the window length theoretically has to be at least 14 seconds, mostly a window length between 30 and 60 seconds is chosen.

As a matter of fact, such window lengths for computing spectral HRV values are lower than indicated by the guidelines of the committee on HRV spectral analysis headed by Berntson that gives the advice to use at least segment durations of 10 times the target rhythm, which means for instance about 100 seconds for spectral measures from the mid frequency band and between 30 and 60 seconds for the high frequency band (Berntson et al., 1997). There are three aspects with regard to these guidelines. The first has to do with reliability of spectral band values, which increases linearly with window length. The second aspect has to do with the reasoning that for obtaining the same reliability, longer time durations are necessary for lower frequencies than for higher ones. This latter reasoning is theoretically not correct, given the assumption that the data are normally distributed (Bendat & Piersol, 1986). At a given segment length the number of degrees of freedom for each component of the power spectral density function is the same for all frequencies. The third is related to findings that HRV increases with increasing segment lengths. This is probably related to the fact that HRV values are not homogeneous distributed over the frequencies. Practically, it is obvious that it is less easy for the lower frequencies to prove that the data follow the required statistical distribution than for higher frequencies. This will probably be the most important reason that it is advised to have longer time segments for lower frequencies.

Three frequency bands have frequently been used: a low (0.02 – 0.06 Hz), mid (0.07 – 0.14 Hz) and a high frequency band (0.15 – 0.40 Hz). The high frequency band mainly reflects respiratory related HRV (RSA), the mid frequency band reflects baroreflex related HRV, while the lower frequencies mainly are related slow task adaptation and temperature regulation (Mulder,1992). The procedure that will be described in the following will be applied to the mid and high frequency band, although there is no specific reason to restrict the method to these bands.

The described spectral profile method works as follows. The heart rate data to be analysed, which may last for several hours, are segmented in small highly overlapping time segments (e.g. window length: 60 seconds; step size: 5 seconds). For each of these time segments the spectral power in a number of specified frequency bands is computed, as well as the mean of the signal (heart rate) in exactly the same manner as is done for longer segments in normal HRV spectral analysis (Mulder, 1992). The time course of the spectral power in each of these specified frequency band is called a spectral profile. In the same manner also spectral profiles can be computed from the coherence, gain and phase functions between cardiovascular time series (e.g. systolic blood pressure-heart rate), again in each of the specified frequency bands (Mulder, 1992).

Theoretically it can be derived that it takes half of the window length before the optimal (theoretical) level of HRV is reached after a sudden change to a new level. In this sense the approach corresponds completely with the output pattern of a moving average filter. Extrapolating this idea it can be concluded that the time resolution of this technique is directly related to the window length; the step size is only relevant for the interdependency of consecutive spectral band values. The smaller the step size, the more smooth the time course of the profile will look. For practical research it is clear that when changes in task demands in consecutive segments have a duration of far less than the time window, the related HRV differences can not be detected. If the duration of such HRV-effects is longer than the time between consecutive task-events, overlap occurs. If there is little overlap, the HRV task-effects will be less than in cases without overlap, but probably differences can be detected. Additionally, the longer the time duration of the segments to be analysed, the higher the reliability of the computed spectral values will be. This reliability is here improved by taking the weighted average of a number of segments belonging to a specified class of task segments.

Study 1: SON intelligence test

The main goal of this study (Rusthoven, 2006) was to find out whether there is a direct relation between task performance in a verbal intelligence test and invested mental effort, as measured with HR and HRV profiles. An adapted version of the SON-R5.5-17 (Snijders-Oomen non-verbal intelligence test for ages between 5.5 and 17 years) was used (Tellegen, 2005). In particular it was studied whether there were indications that children were giving up in cases that the test items could not be solved adequately.

Methods

A specific task was constructed using four subtests of the SON-R5.5-17 (further indicated as the SON). Thirty-seven children in the age category between 9 and 11 years, with normal education and without known developmental problems, were tested. The four selected subtests were, respectively: Categories, Analogies, Mosaics and Patterns. The first two are abstract reasoning tests; the last two require mainly spatial insight (Snijders, Tellegen, & Laros, 1988). Each subtest consisted of two series of items, each with increasing level of difficulty. The number of items per subtest varied with the required solution time, with Categories and Analogies being the shortest (about 15 seconds) and Mosaics and Patterns being the longest (about 70 seconds). The short subtests had a maximum of 10 items per series, while the long series included 7 items as a maximum. A series of items was ended when 3 errors were made in the Patterns and Mosaics subtests and 5 errors in the Categories and Analogies subtests. Each new series started with easy items. In order to prevent effects of time-on-task, half of the subjects started with the short lasting items, while the other half started with the items of longer duration. Moreover, after the first and after the third subtest a resting period of 3.5 minutes was included that was used as baseline measurement.

HR and HRV measures were obtained from hardware triggered R-peak events in the ECG and were corrected for artefacts, using a combination of automatic detection/correction and visual inspection followed by manual correction where necessary in the Carspan spectral analysis programme (Mulder, 1992). Spectral profiles from HR and HRV (mid and high frequency band) were computed with the same programme. Data were analysed with a window length of 30 seconds and a time step of 1 second. With respect to the subtests, around each item a small time segment was selected that was aligned around the response of the participant. For Categories and Analogies this window started 20 seconds before and ended 6 seconds after the response. For the subtests with longer duration this segment was between 80 seconds before and 10 seconds after the response (for Patterns) or between 70 seconds before and 10 seconds after the response (for Mosaics). Two types of differences were studied: between the first and second half of each series and between items with correct and incorrect answers. In this way 4 averaging categories were obtained: CB: correct items in the first half (begin); CE: correct items in the second half (end); IB: incorrect, first half; IE: incorrect second half. It has to be realised that these categories are not completely independent: in the first half of each series more correct answers may be expected because the items are easier. Moreover, the defined time segments were further adapted by shortening these by the individual mean durations of the items in the CB, CE, IB and IE categories. Finally, HR and HRV profile-values were averaged over these newly defined segment values. For reasons of a having a more normal distribution, the natural logarithm of the HRV measures was taken (Van Roon, 1998).

It was expected that HR is higher and HRV lower in the more difficult (effortful) items; items in the first half (CB, IB) are estimated as easier than those in the second half (CE, IE), while correct answered items (CB, CE) are estimated as easier than incorrect items (IB, IE). No signs of giving up during difficult items were expected. HRV was expected to be higher in the baseline conditions, while HR would be lower in the two (averaged) baseline conditions. Data were statistically analysed using twosided paired t-tests after Bonferoni-Holm correction (Van Peet, Wittenboer, den Hox, 2001). For all tests a significance level of 5% was applied.

Results and discussion

Only the results of the subtests Categories (figure 1) and Patterns (figure 2) will be described in detail. From other subtests only general effects will be mentioned.

In all subtests HRV (mid and high frequency bands) is lower during task than in the baseline conditions; HR is higher during task performance than in the baseline conditions, for the tests with the longer lasting items (Patterns, Mosaics), but not for the subtests Categories and Analogies. HRV-high distinguishes items from the first and the second half in all the subtests except Mosaics: HRV is higher in the first half; the same pattern is found when comparing correct with incorrect items, but is only significant in the Categories and Patterns subtests. HRV-mid shows similar patterns as HRV-high, but the level of significance is only reached in the Patterns subtest (both comparing first-second half and correct-incorrect items). The Mosaics subtest shows a different pattern: no differences are found on HRV (high and mid frequency band) between first and second half and correct versus incorrect items, although all task conditions are strongly different from the baseline condition in the expected direction. It is remarkable in this case that HR is different between first-second half and correct-incorrect items, but in the other direction than expected: HR is higher in the first half than the second and higher during the correct items. As a matter of fact this same pattern was found for HR in the other subtests, but not significantly.

In more detail: figure 1 shows the results for the Categories subtest. Although HRV-high and mid frequency bands show very similar patterns, the effects are only significant for the high frequency band. Note the small (non-significant) differences on HR between baseline and tasks and the task conditions mutual.

Figure 2 shows the results of the Patterns subtest. Now all HRV effects are clearly significant, both when comparing task with baseline values as well as the task conditions mutual. HR is higher during task performance than during baseline conditions, while the task conditions do not differ mutually.

The study shows that there are signs of increased effort in incorrect items, therefore, no arguments are found for the idea that children give up during the most difficult items. All together, the results are remarkable clear and consistent, while interesting details are found between the subtests and the used variables. In general, HRV-high seems to be the most sensitive variable; this holds in particular for the subtests with the shortest item-durations (Categories, Analogies). For the Pattern subtest high and mid frequency band have similar sensitivity. For Mosaics a quite different pattern occurs in both bands: no differences between the task conditions, but large differences with baseline values, probably related to the high difficulty level of all items.

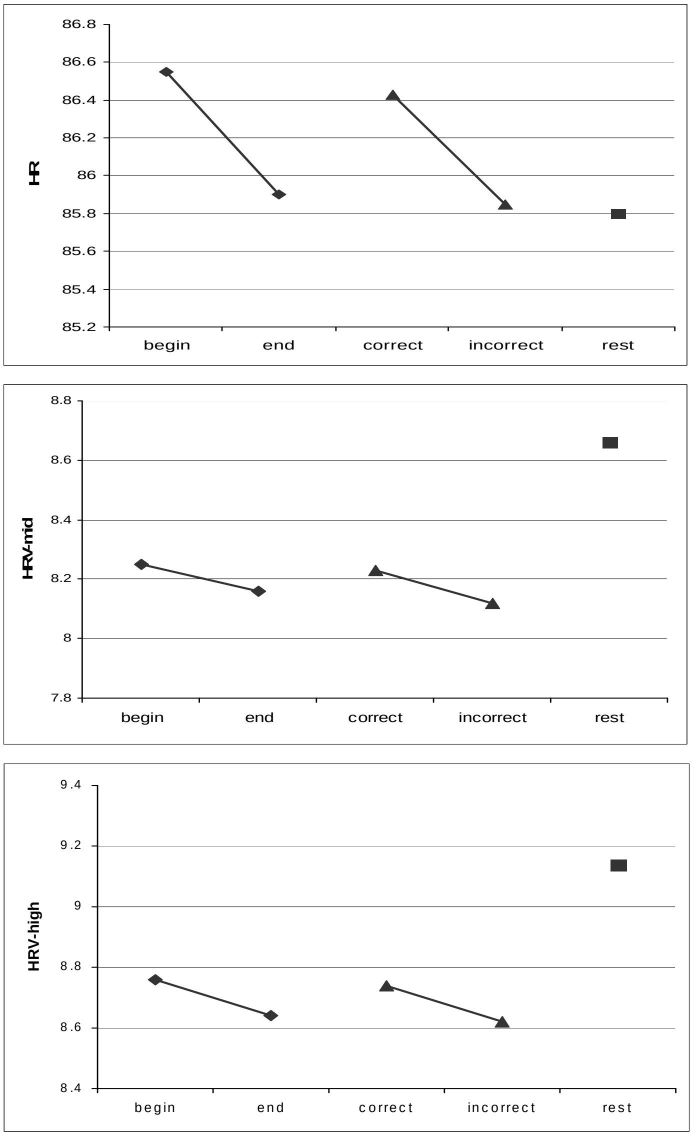

Figure 1. HR and HRV results of Categories sub-test. HR (upper panel), HRV in mid (middle panel) and high frequency band (lower panel) as a function of task conditions. Begin/end indicates first/second half of the items, correct/incorrect indicates correct vs. incorrect items, rest indicates baseline data. HRV measures are natural logarithms

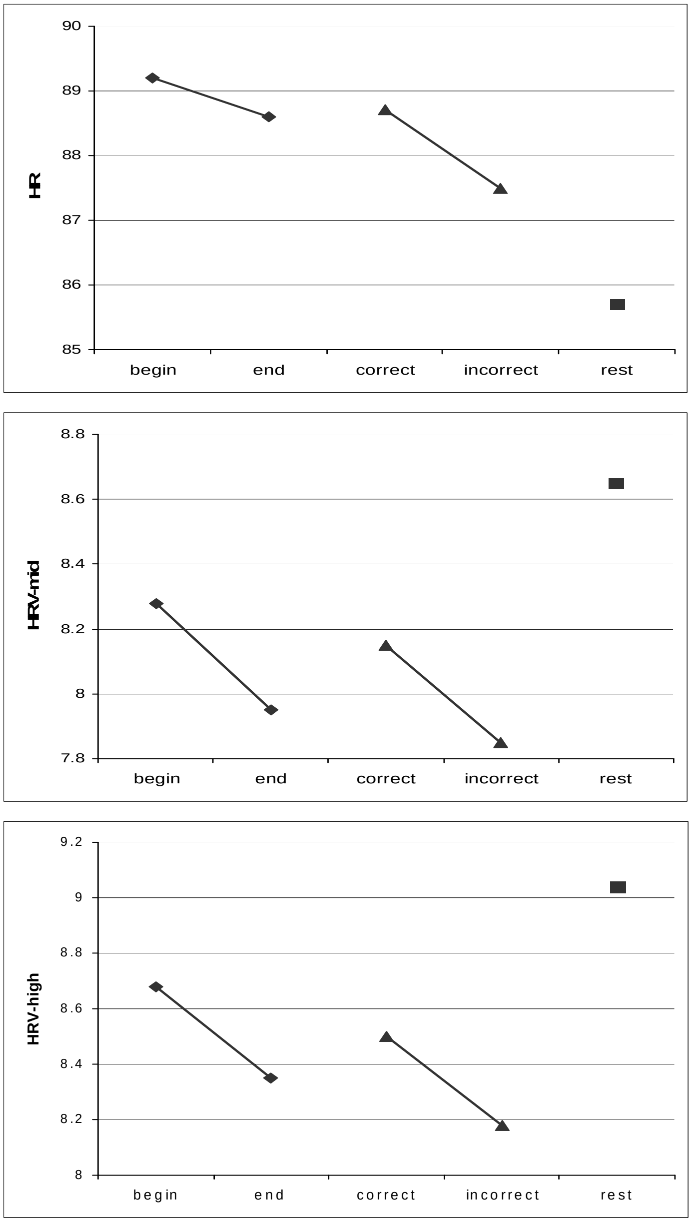

Figure 2. HR and HRV results of Patterns sub-test. HR (upper panel), HRV in mid (middle panel) and high frequency band (lower panel) as a function of task conditions. Begin/end indicates first/second half of the items, correct/incorrect indicates correct vs. incorrect items, rest indicates baseline data. HRV measures are natural logarithms

So, in contrast to many other laboratory studies, it seems that here the high frequency band is more sensitive to changes in mental effort than the mid frequency band in the described setup for the subtests with the short lasting items. This finding corresponds well with the position of the Committee Report in Psychophysiology (Berntson et al., 1997) that segments of longer duration are necessary for lower frequency bands. Whether this has to do with the statistical distributions of these variables or with the fact that there is more overlap in HRV profile-patterns between the segments is still unknown.

In many laboratory experiments with heavy mental loading tasks (memory search and counting, mental arithmetic, etc.) it is found that HRV in the mid frequency band is reduced with about a factor 2 (for natural logarithm this is a difference of 0.7), while differences have to be about 0.4 (on a ln scale) for getting significant results, assuming about 20 participants (Van Roon, 1998). In the present experiment with 37 participants, differences in the high frequency band lower than 0.2 were significant (Categories, Analogies). This means that the presented method is certainly not less sensitive than when having united task segments.

Conclusions

- There are signs of increased effort when dealing with the most difficult items of the test.
- Despite the short duration of the items stable HRV differences can be found.
- The high frequency band of HRV seems to be more sensitive in the tests with the shortest item durations than HRV the mid frequency band, probably related to the restricted time-resolution of the latter.
- Comparison of these results with other types of task can probably clarify whether these findings are task related or have to do with the item duration itself.

Study 2: mental effort in simulated flight

The main aim of this study (De Rivecourt & Kuperus, 2006) is to find out whether different task difficulty levels of a simulated flight are reflected in corresponding HR and HRV profiles. Knowing that in this task the pattern of task demands is changing very rapidly one has to find out whether it is still possible to work with HRV profiles that require relative long time durations. Moreover, it is has to be found out which variables are the most sensitive for this approach.

Method

Nineteen male participants from the Royal Netherlands Air Force selection and training programme, in the range of 17 to 27 years, performed an instrument flight task on an ALSIM AL 100 Flight Trainer. The participants had some experience in a flight simulator but had not yet started their military flight education. They all had less than 100 flight hours and can be considered as relatively inexperienced.

After a short instruction and familiarisation in a training flight (24 minutes) participants were expected to have sufficient control and track of the task to be performed. Before the experimental flight (that lasted 28-30 minutes) they were instructed about the flight profile, which was available during the whole flight in a written form attached to their knee. The participants had to perform an instrument flying task with increasing task complexity, starting with straight and level flight, followed by a horizontal turn, a climbing turn, a descending turn and combinations of these elements with at the end relatively fast transitions. Six flight segments could be distinguished, which lasted each between 1.5 and 2.5 minutes. More detailed task analysis led to a description of the task in 39 consecutive flight elements that could be assigned to four increasing levels of task complexity: 1. horizontal flight, 2. horizontal turn or climbing/descending, 3. turns in combination with climbing/descending, and 4. transitions. Transitions from one phase to another were considered to be more complex than category 3. The duration of each of these flight elements was variable and lasted between 5 and 100 seconds. In particular, most transitions were not longer than 20 to 25 seconds. Although the flight scenario was very detailed and precisely prescribed in time, individual variations could occur when participants started a new phase too early or too late.

HR and HRV profiles were obtained in the same way as described in study 1, while in this case, for reasons not relevant for the present chapter, the window length for HR was 4 seconds, while it was 30 seconds for the HRV profiles. The step size was 1 second for all variables. HRV power values for the 4 levels of complexity were obtained by computing weighted averages of the short segment HRV profile values. Baseline values for the heart rate data were computed by taking the average of a five-minute rest before the start of the flight and a same resting period after the flight. Next to heart rate also eye fixation patterns were measured, as well as subjective indications of task complexity in the 6 consecutive flight phases. These data will be described elsewhere.

To summarize, data are analysed in two ways: on the basis of the 6 consecutive flight segments and on the basis of the segments belonging to the 4 levels of complexity. It is expected that HR increases and HRV deceases with increasing task demands. This will be tested using multilevel analysis (MLWin) with a random effect model (Wright, 2004; De Rivecourt & Kuperus, 2006). A one-sided significance level of 5% was used for all tests.

Results and discussion

The HR and HRV results with respect to the 6 consecutive flight segments are illustrated in figure 3. In comparison to the baseline values HR is strongly increased in all flight segments, while there is a continuous ongoing increase of HR as a function of time (and task load, as task load increases as a function of flight time). Multilevel analysis shows that all segments differ from rest, while segments 1, 2, 3 and 4 can be distinguished from each other. No differences are found between segments 4, 5 and 6. The HRV patterns in both the mid and high frequency band show differences between the first segment and all others, while these other segments are mutually not different, except segment 6 for the high frequency band which is statistically different from the other segments. Segment 1 does differ from baseline in both frequency bands.

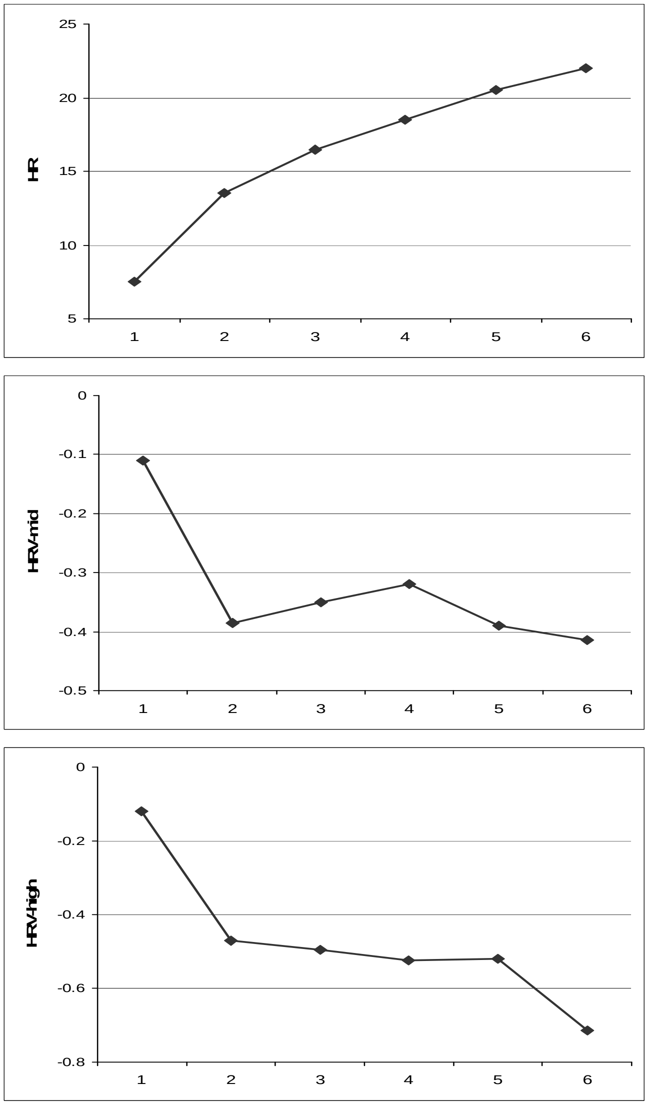

Figure 3. HR and HRV results as a function of flight segment. Changes in HR (upper panel), HRV in mid (middle panel) and high frequency band (lower panel) as a function of flight segment complexity; level 1 corresponds with the most simple task elements, level 6 with most complex ones. All values are related to baseline. HRV measures are natural logarithms

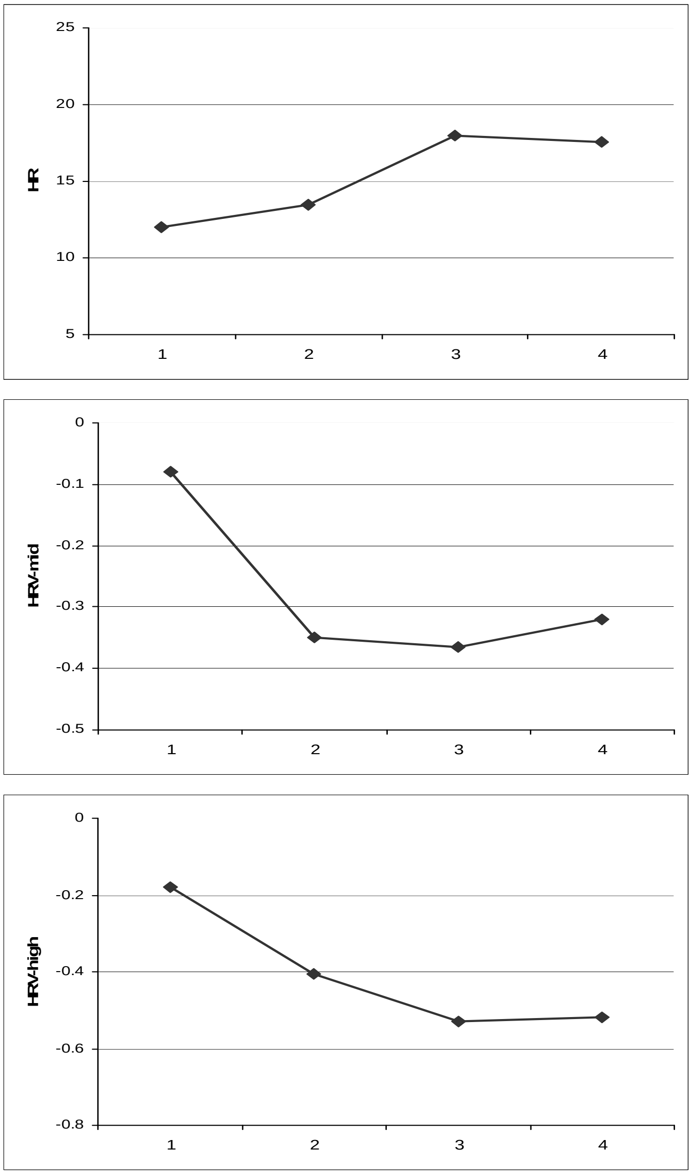

Figure 4. HR and HRV results as a function of task levelChanges in HR (upper panel), HRV in mid (middle panel) and high frequency band (lower panel) as a function of increasing task level complexity: 1. horizontal flight, 2. horizontal turn or climbing/descending, 3. turns in combination with climbing/descending, and 4. transitions.. All values are related to baseline. HRV measures are natural logarithms

The HR and HRV results on the basis of the four task demand levels are illustrated in figure 4. For HR, all task demand levels differ from the baseline, while HR increases with task demand level. Multilevel analysis shows that level-pairs 1,2 differ from level-pairs 3,4, while the levels are not mutually different within these pairs. HRV decreases strongly with increasing task demand level in both frequency bands. Multilevel analysis shows that in both frequency bands level 1 can not be distinguished from the baseline periods. In the mid frequency band levels 2, 3 and 4 differ from level 1, while these levels do not differ mutually. In the high frequency band level 1 differs from level 2, level 2 differs from the pair 3,4, while these latter levels can not be distinguished mutually.

The HR data in this study show much larger increases compared to baseline values than in Study 1. This finding corresponds with other studies in flight simulators and in driving simulators, which illustrates strong effects of manual control on the HR level in comparison to mental tasks that require more working memory operations (Veltman & Gaillard, 1996, Wilson, 1992, De Waard, 2002). HRV level changes in the present study are in the same range as those found in Study 1 and in many other mental loading studies (Van Roon, 1998). Although there seems to be a problem in this study that task load effects are confounded by time-on-task, it has to be mentioned that in general time of task effects are in the opposite direction than found in this study. Usually HR decreases, while HRV increases as a function of time on task (Mulder, et al. 2003).

The two approaches, one based on continuous task segments with lengths between 1.5 and 2.5 minutes and the other based on short segment analysis related to task demand level give, in great lines, similar results. For HR, the analysis based on consecutive segments seems to be the most informative, because more levels can be distinguished. For HRV-high the task demand level analysis seems to be at least even informative as the segment analysis, because in this case the levels 2 (curves or climbing/descending) and 3 (curves in combination with climbing/descending) can be distinguished. Furthermore, using this task demand level analysis the problem of untangling time on task effects from task load effects is less prominent. Moreover, in the present setup differences in the high frequency band of HRV of about 0.15 (ln scale) can be distinguished, both in the segment analysis (level 6 compared to 2, 3, 4 and 5) as in the task demand analysis (level 2 compared to 3), which is quite sensitive compared to other studies using data segments of longer duration.

Conclusions

- Both HR and HRV differences can be detected between the different task segments as well as between task demand levels.
- The differences in HR, compared to baseline are remarkably larger than in Study 1.
- Again, HRV in the high frequency band seems to be more sensitive in both types of analysis than HRV in the mid frequency band.
- Also in this study the short segment analysis based on the profiles approach is suitable, both for HR and HRV analysis.

- This study shows, once again, that the combination of HR and HRV measures gives more information than taking just one of these types of measures.

General discussion and conclusion

The proposed short segment analysis technique based on (spectral) profiles worked well in both experiments. Despite the short duration of some segments (certainly in Study 2, related to flight transitions) different task demand levels could be distinguished by averaging HRV band values and HR values for relevant short segments. It is certainly not denied that overlap will exist between the HRV effects in consecutive short task segments, but the results of these two short studies show that despite this overlap relevant task effects can be found at the level to be studied. In general, it may be expected that the less overlap exists between segments the better the statistical resolution will be. This is a hopeful conclusion for all those experimental situations at which partly overlap can not be avoided. Moreover, it is remarkable that the standard error of the spectral band values (not shown in the data presentation) is not larger for baseline data (relative long consecutive segments) than for short segments from, for instance, flight transitions.

In both studies the HRV effects in the mid and high frequency band showed about the same pattern, but in both cases the effects in the high frequency band were more sensitive to task load. Effortful laboratory tasks in the past (Van Roon, 1988) showed more consistent effects in the mid frequency band, based on longer lasting consecutive task durations (about 5 minutes of data). At least two explanations can be given for this difference. The first is related to possible statistical distributions of these spectral band data. The 10-second rhythm constituting the power in the mid frequency band shows in general, for instance during longer lasting baseline measurements, an on-off going burst pattern. If this occurs in the same way during the type of tasks as described in this paper, large random variations may be expected; these will then result in large standard errors and statistical distributions far from normal. The second possible reason has to do with control of the autonomic nervous system. It may be expected that HRV patterns in the high frequency band are directly related to vagal control, while changes in the mid frequency band are related to both vagal and sympathetic effects (van Roon, 1998). Knowing that in sympathetic control much longer time constants are involved, having effects over 10 or 20 seconds, while in comparison vagal effects are restricted to 2 or 3 seconds it is to be expected that much more overlap will exist in effects in the mid frequency band than in the high frequency band. This would then result in smaller differences between different types of task segments in the mid frequency band compared to the high frequency band.

It can be concluded that the described short segment spectral profile approach looks very promising for use in (semi-)realistic working situations. The important pre-requisite is that on the basis of simple task analysis relevant segments have to be detected that do not overlap too much with respect of the HRV patterns. It is advised to use both HR and HRV measures (in both the high and mid frequency band); combination of these measures gives more information than taking just one type of measures.

Acknowledgments

The authors wish to thank the staff of the Centre for Human Factors in Aviation (Soesterberg, The Netherlands) for their support, the Air Force cadets, as well as Evelien Kums, Michiel de Leeuw, and Siemen Henk Sikkema for fruitful co-operation in Study 2. In Study 1 the help of parents and participating children and the intensive support of Peter Tellegen and Bas Kortmann are greatly acknowledged. For both studies we thank the technical staff of the department of Psychology of the University of Groningen for their support.

References

Bendat, J.S. & Piersol, A.G. (1986). *Random data: analysis and measurement procedures*. New York: John Wiley and Sons.

Berntson, G.G., Thomas Bigger JR, J., Eckberg, D.J., Grossman, P., Kaufmann, P.G., Malik, M., Nagaraja, H.N., Porges, J.P.S., Stone, P.H. and Van der Molen, M.W. (1997). Committee Report - Heart rate variability: Origins, methods and interpretive caveats. *Psychophysiology, 34*, 623-648.

Boucsein, W. & Backs, R.W. (2000). Engineering Psychophysiology as a Discipline: Historical and Theoretical Aspects. In R.W. Backs and W. Boucsein (Eds.). *Engineering Psychophysiology: Issues and applications* (pp. 3-30). Mahwah, NJ: Lawrence Erlbaum.

De Rivecourt, M., & Kuperus, M.N. (2006). *Heart rate and eye activity measures as indices for mental effort in aviation.* Master thesis. Groningen: University of Groningen.

De Waard, D. (2002). Mental Workload. In R. Fuller and J.A. Santos (Eds.) *Human Factors for Highway Engineers* (pp. 161-175). Oxford: Pergamon.

Mulder, G. (1980). *The heart of mental effort*. Ph.D. thesis. Groningen: University of Groningen

Mulder, G. (1986). The concept and measurement of mental effort. In G.R.J. Hockey, A.W.K. Gaillard, and M.G.H. Coles (Eds.), *Energetics and human information processing* (pp. 175-188). Dordrecht, The Netherlands: M. Nijhoff Publishers.

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

Mulder, L.J.M., Van Roon, A., Veldman, J.B.P., Laumann, K., Burov, A., Quispel, L., & Hoogeboom, P. (2003). How to use cardiovascular state changes in adaptive automation. In G.R.J. Hockey, A.W.K. Gaillard, and O. Burov (Eds.), *Operator Functional State*. (pp. 260-269). Amsterdam: IOS Press

Nickel, P. & Nachreiner, F. (2003). Sensitivity and diagnosticity of the 0.1 Hz component of heart rate variability as an indicator of mental workload. *Human Factors, 45*, 575-590.

Rusthoven, H. (2006). *De hartritmevariabiliteit tijdens afname van de SON-test. Masters thesis.* Groningen: University of Groningen.

Snijders, J.Th., Tellegen, P.J. & Laros, J.A.(1988). SON-R 5.5-17 *Verantwoording en handleiding.* Groningen, The Netrherlands: Wolters-Noordhoff BV.

Van Steenis, H.G. (2002). *On Time Frequency Analysis of Heart rate Variability.* PhD thesis. Rotterdam, The Netherlands: Erasmus University.

Van Peet, A.A.J., Wittenboer, G.L.H., & Van de Hox, J.J. (2001). Toegepaste statistiek: Inductieve technieken. Groningen, The Netherlands: Wolters-Noordhoff BV.

Van Roon, A.M. (1998) *Short-term cardiovascular effects of mental tasks, physiology, experiments and computer simulations*. PhD thesis. Groningen, The Netherlands:University of Groningen.

Veltman, J.A. & Gaillard, A.W.K., 1996, Physiological indices of workload in a simulated flight task. *Biological Psychology*, *42*, 323-342.

Wilson, G.F. (1992), Applied use of cardiac and respiration measures: practical considerations and precautions. *Biological Psychology*, 34, pp. 163-178.

Wright, D.B. (2004). Statistical software review. *British Journal of Mathematical & Statistical Psychology*, 57, 383-384.