

# Heart Rate Variability: Indicator of User State as an Aid to Human-Computer Interaction

**Dennis W. Rowe**

The MITRE Corporation  
1820 Dolly Madison Boulevard  
McLean, VA 22101 USA  
+1 703 883 5883  
drowe@mitre.org

**John Sibert**

The George Washington  
University  
Washington, D.C. 20052 USA  
+1 202 994 4953  
sibert@seas.gwu.edu

**Don Irwin**

MAJ USAF HQ USAF/TEP  
1530 Air Force Pentagon  
Washington, D.C. 20330 USA  
+1 703 695 0900  
irwind@seas.gwu.edu

## ABSTRACT

This preliminary study explores the use of Heart Rate Variability (HRV) as an indicator of user state. In the study, a visual display is used to vary the levels of a complexity factor to assess the impact on user mental effort in a monitoring task. Mental effort is measured both subjectively and physiologically. Two findings indicate the potential value of HRV. First, user HRV showed significant discriminatory sensitivity to the manipulation of the independent variable on the basis of domain experience. Second, HRV appeared to indicate the point at which user capacity to process targets was exceeded. Results warrant further investigation but suggest the use of HRV as a tool for design and analysis of user interfaces, and as a research basis for use as component of an adaptive interface.

## Keywords

Heart rate variability, mental effort, user state, air traffic management, airspace complexity, free flight

## INTRODUCTION

Over the past three decades, human monitoring of visual displays has become relatively commonplace. During this same timeframe however, the information to be monitored has become increasingly more complex, often in both pattern and in sheer volume of information displayed. Because of this increased complexity, situations requiring human monitoring are apt to create cognitive loads on users that can adversely affect their performance. Errors resulting from cognitively burdened users may have disastrous consequences in terms of monetary losses, or worse, the loss of human lives [1].

We believe that having a computer monitor the state of the user via physiological measures can alleviate challenges brought about by the complexity of information on the visual display. Evolving computer technology should

enable interfaces to be developed that possess a special sensitivity to the states of attention and intentions of users as suggested by Velichkovsky and Hansen [19]. Computer knowledge of user state should not only be useful in designing effective interfaces, but also in adapting interfaces to accommodate individual user differences.

It is widely recognized that perceptions of complexity vary from individual to individual. Differences among individuals such as age, experience, and training, to name a few, account for some of the variation in these perceptions [11]. The important role that individual differences play in the way humans perceive the visual display information highlights the need for designing future human-computer interfaces that are sensitive to the capabilities and limitations of individual users. Computer system designs that account for individual differences will be of value, as they will foster a human-machine synergy allowing both the human and computer system to operate at peak efficiency. Given their unique "signatures" and often high correlation with perceived workload, physiological measures of mental effort are perhaps the most likely areas to draw from in order to provide a foundation for the analysis and design of future user interfaces [17].

## *Advantages of Physiological Measures in Determining Mental Effort*

The advantages of using physiological measures to determine mental effort are manifold. First, they are relatively unobtrusive; assuming of course that users adapt to the few transducers that are affixed to their bodies. Secondly, they provide measures which do not require overt performance; that is, most physiological measures can be recorded in the absence of overt behavior. It should be noted, however, that it is often advantageous to obtain measures of both performance and physiology in order to infer changes in user strategies and workload with variations in system demands. Third, physiological measures are inherently multi-dimensional and can provide a number of 'views' of user mental workload.

Permission to make digital/hard copies of all or part of this material for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copyright is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires specific permission and/or fee.

CHI 98 Los Angeles CA USA  
Copyright 1998 0-89791-975-0/98/ 4..\$5.00

Lastly, since most physiological signals are recorded continuously, they offer the potential for providing measures that respond relatively quickly to phasic shifts in mental effort and workload [8].

Several physiological measures have been used as an index of user mental effort and mental workload. Included are two general classes of physiological measures: central nervous system and peripheral nervous system measures. Central nervous system measures include electroencephalographic (EEG) activity, event-related brain potentials (ERP), magnetoencephalographic activity (MEG), measures of brain metabolism such as positron emission tomography (PET), and measures of electro-oculographic (EOG) activity. Peripheral nervous system measures include cardiovascular measures, measures of pupil diameter, respiratory measures, and measures of electrodermal activity (EDA) [8]. Each of these measures has its strengths and weaknesses in terms of the relative difficulty in collecting, analyzing, and interpreting their meaning; each also differs along the lines of sensitivity, diagnosticity, intrusiveness, reliability, and generality of application.

The most popular physiological techniques employed in the assessment of mental effort and mental workload in the last 30 years have been measures of cardiac activity [21]. Sensitivity of a number of different cardiac measures to variations in workload have been examined. These techniques include: the electrocardiogram (ECG), blood pressure measures, and measures of blood volume. Of these three, measures of electrocardiographic activity have shown the most promise [8].

In the current study, we examined several possible measures for assessing user state [21]. Based upon the criteria of sensitivity, diagnosticity, ease of data collection, and nonintrusiveness, we decided to use a measure of electrocardiographic activity; specifically, Heart rate variability (HRV).

This study examines the efficacy of using HRV under conditions of varying complexity to provide feedback to a computer on the state of the human user. We envision that the computer system can use this feedback to form the basis for adapting to individual user differences, thus creating a human-computer system synergy for ensuring acceptable levels of human performance.

The study seeks to answer the following broad question: To what extent can HRV be used as an indicator of user effort under conditions of varying workload? This question has the following sub-components:

- How do changes in the complexity factor affect subjective measures of user mental effort and user workload?

- How sensitive is HRV to changes in the complexity factor? What is the nature of this sensitivity?
- What is the relationship between HRV and the subjective measure of user mental effort? To what extent, and under what conditions do these measures correlate?

We believe that answers to these questions can support the use of HRV as a means of identifying problem areas of user interfaces. We also think that answers to these questions may point to the use of HRV as real-time mechanism for adapting visual displays to the needs of individual users.

### Background

#### *Acceptance of HRV as a Measure of Mental Effort*

Since the early 1980s, Heart rate variability (HRV), or sinus arrhythmia, has gained widespread acceptance as a measure of mental effort. HRV is a measure of the oscillation of the interval between consecutive heartbeats. The literature is replete with references to this measure that are expressed in terms of consecutive cardiac cycles. Such terms include: cycle length variability, heart period variability, R-R interval variability (referring to the beat-to-beat interval formed by consecutive R-waves of the cardiac signal), and the R-R interval tachogram [6, 7].

The use of HRV in both laboratory and field settings is valued not only because of its nonintrusiveness, but also because of its utility where continuous recording is required [17]. In laboratory studies, HRV has consistently responded to changes from rest to task conditions and to a range of between-task manipulations [1,13,16]. In operational contexts, HRV has seen increased use as an indicator of the extent of task engagement in information processing requiring significant mental effort, particularly in flight-related studies [8, 16, 17, 22, 23]. HRV has also responded rapidly to changes in user workload and strategies, usually within seconds [1, 2]. Thus, HRV has been able to detect rapid transient shifts in mental workload [8].

#### *Spectral Analysis of Sinus Arrhythmia*

Spectral analysis methods have been applied to the cardiac interval tachogram. Power spectral density analysis provides information on how power (or variance) distributes as a function of frequency. With the cardiac interval tachogram, the power spectral density analysis yields a measure of how the heart fluctuates as a result of changes in the autonomic nervous system [7].

Spectral components of the cardiac interval signal for short-term changes in HRV fall into three frequency ranges: the Very Low Frequency (VLF) range, the Low Frequency (LF) range, and the High Frequency (HF) range. For long-term changes in HRV, e.g. 24 hour recordings, the Ultra Low Frequency (ULF) range is also examined [6].

The LF component, known as the baroreflex band, is of chief interest as an indicator of cognitive workload. This component, commonly referred to as the 0.10 Hz component, reflects short-term changes in blood pressure and spans a range from 0.04-0.15 Hz. That is, there are approximately 6 fluctuations of the heart per minute due to changes in blood pressure. A peak in this component is indicative of lower cognitive workload conditions. A flattening of this component reflects conditions of greater mental workload [7].

#### *Sparse Discrete Fourier Transform (SDFT)*

A method of computing Fourier transforms on the cardiac interval signal which requires relatively few data points for computation is known as the Sparse Discrete Fourier Transform (SDFT) [12]. This method is ideal for recordings of short duration where 300 to 500 data points exist for computing spectral values. With SDFT, the times between successive R-peak signals from selected overlapping groups of R-peak waves are used to compute spectral values for the 0.10 Hz component of the cardiac interval signal. Thus, SDFT computations provide a trace of the change in the 0.10 Hz component of the cardiac interval signal over time so that changes in mental effort can be recorded throughout the entire period of a trial.

#### *Use of HRV in Determining Mental Effort*

In a recent study involving the level of user control and changes in HRV during simulated flight maintenance, the demands of dynamic monitoring and fault diagnosis for 11 trainee flight engineers were examined in relation to changes in HRV [17]. HRV was found to be sensitive to the different phases of the work environment. In particular, the frequency 0.07-0.14 Hz range was suppressed during the mentally demanding problem solving mode. The findings of this study support both the use of HRV as a physiological index of mental effort and its value in operational contexts.

In a 1987 study, Aasman, Mulder, and Mulder found HRV to be associated with the changing levels of user effort. In this study, participants were given a simple (non-counting) and complex (counting) condition of a task. In the simple condition, participants were required to indicate the presence or absence of a target by pressing one of two buttons. In the complex condition, participants were required to keep a running mental count of each memory set item. In the complex condition, the operations of updating the counters and rehearsing them had to be time shared with other operations. Participants rated this task as one requiring great mental effort.

The 1987 study showed that the amplitude of the 0.10 Hz component of the cardiac interval signal was particularly affected in the complex task condition, as long as the subjects were working within the limits of working memory. When the limits of working memory were

exceeded, most subjects were unable to cope with the demands of the task as evidenced by a performance decrement and an increase in HRV. Thus, when working memory was exceeded, participants gave up, indicating that less effort was invested.

Mental effort expended by a user has not necessarily been found to be related to task load [20]. The amount of effort invested by a user is determined by internal goals and criteria they choose to adopt. Situations have been observed in which an increase in task demands resulted in a decrease in mental effort [18]. The explanation for this seemingly paradoxical result is that the task became so difficult that performance could not be maintained at an acceptable level so users settled for lower performance and reduced the amount of mental effort they were investing in the task.

In our recent informal study, which provided the foundation for the current study, the sensitivity of HRV to the cognitive demands of playing different versions of a computer game was examined. Participants included 2 males and 4 females who ranged in age from 11 to 19 years. The game was an abstraction of an air traffic control scenario in which some objects moved across the screen following fixed routes, while others moved in "off route" patterns. Participants were asked to keep "like" objects from colliding as they moved across the screen.

The primary goal of this experiment was to answer the following question: How is HRV affected by the presence of certain combinations of intersecting routes and "off route" objects, both in the time and frequency domains. Results of this experiment indicated that there was no statistically significant effect of the number of intersecting routes on HRV,  $F(2, 10) = 0.22, p=0.81$ . With the number of "off route" objects, however, there was an effect at the 0.15 level,  $F(2, 10) = 2.31, p= 0.15$ .

#### *A taxonomy of display monitoring tasks*

Display monitoring tasks fall into two broad categories: those involving monitoring of relatively static displays and those involving monitoring displays with movement. Static monitoring tasks include those where information appears in the same relative position, but values are subject to change. Examples include digital readout displays and displays with dials, knobs or scales whose indicators change value in order to reflect changes in system status.

Monitoring tasks involving movement offer a variety of combinations of different parameters. In general, tasks with movement may be classified as those with constrained movement and those with unconstrained movement. In tasks with unconstrained movement, objects may move on the display in any direction and at varying speeds. In addition, there may also be encoding associated with objects which indicates movement in more

than one direction. Classic examples include 3-D and 4-D air-to-air and air-to-ground intercept applications, as well as many air traffic management applications.

Monitoring tasks involving constrained movement are those in which an object's location, direction, or speed is moderated or bounded in some fashion. That is, objects may be confined to movement along predetermined paths, or with a specified velocity or a limited range of velocities. As with unconstrained movement tasks, there may also be encoding which indicates movement in more than one dimension.

Within the overall task space of display monitoring, we chose a task for the current study which required monitoring of moving, relatively constrained objects. In particular, this task approximated that employed in air traffic management using an air traffic management visual display. We believe that that the air traffic management visual display offers a rich environment for examining the applicability of HRV as an index of mental effort, in that multiple moving objects are displayed whose parameters can be varied to create different cognitive demands on human monitors.

#### *Air Traffic Management: Decision Making Model*

A simple decision making model documented in an air traffic management training manual and used in the work by Mogford, et al., assumes that a user makes decisions in four phases: scanning, projecting, planning, and acting [9].

In the *scanning* phase, user attention is continuously switched between different stimulus sources. The research team in [9] assumed that the purpose of scanning is to update the mental representation of the visual display and attendant communications systems as the basis for detecting potential problems. As a potential conflict is detected, a user then *projects* the future positions of objects and assesses the likelihood of a conflict. During the *planning* phase, the user reviews possible solutions to the problem. Finally, during the *acting* phase, the user selects and implements a solution. This Scan-Project-Plan-Act paradigm, matched air traffic controller's reports of their decision making processes and was sufficient to guide preliminary research [9].

#### *Air Traffic Management: Complexity factors*

Grossberg defined airspace complexity as, "a construct referring to the characteristics, dynamic and static, affecting the rate at which workload increases", and also defined 12 "complexity factors" which could be used to describe an airspace [11]. Based on the research of Grossberg and Mogford [9, 10, 11], together with a consideration for the most salient factors affecting airspace complexity in the future air traffic management system, we chose two factors for use in the current study: the number of intersecting flight paths formed by fixed

routes of flight, and the number of complex aircraft routings. Based on the results of a pilot study, the number of intersecting flight paths formed by fixed routes of flight was set at 3 as a condition of the experiment. The number of complex aircraft routings, or "Free Flyers", was varied at levels of 0, 4, 8, 12, and 16.

## METHOD

### *Participants*

Thirteen people (age 21 to 60 years) participated in the experiment. All participants had a college degree, were adept at using a mouse, and used a computer daily. Three participants were female. Five participants had Air Traffic Control experience. Each participant was paid for participating in the study.

### *Apparatus*

The experimental task was performed using a Compaq laptop computer with a 120 MHz Pentium processor running Windows 95. The laptop's keyboard was used for text input on the demographic survey, but a mouse was used for all other inputs. A 17" color SVGA monitor was used for display during the game with a resolution of 640x480 and 256 colors.

Electrocardiogram (ECG) equipment consisted of Dataq's WinDaq/200 hardware and software used in conjunction with a Windows 3.1 based 80386 PC [3]. Delphi Developer 2.0 was used to develop the computer game used in the experimental task. CARSPAN and SAS software packages were used for data reduction and analysis.

### *Task*

Participants were required to play a series of five games while connected via three leads to ECG recording equipment. (One lead was positioned near the heart, another on the left wrist, and the third above the right ankle). The games were abstractions of an air traffic management environment designed to simulate a future "Free Flight" paradigm; that is, some aircraft targets moved across the screen following fixed routes, while others followed off-route or "Free Flight" paths. Task difficulty was manipulated by varying the number of free flight aircraft targets. Five games were played, a control game with no free flyers and four games with treatments of 4, 8, 12, and 16 free flyers. A view of the screen with the 16 Free Flyer treatment is shown in Figure 1.

During each game, participants were required to keep objects at the same altitude from colliding as they moved across the screen. Aircraft targets had one of two altitudes: Flight Level 350 (35,000 feet), or Flight Level 330 (33,000 feet). Participants could change altitudes of an aircraft target by using the mouse to click on either the aircraft target symbol "X", or its altitude "350" or "330". Clicking on the 350 altitude changed it to a 330 altitude,

and clicking on the 330 altitude changed it to a 350 altitude.

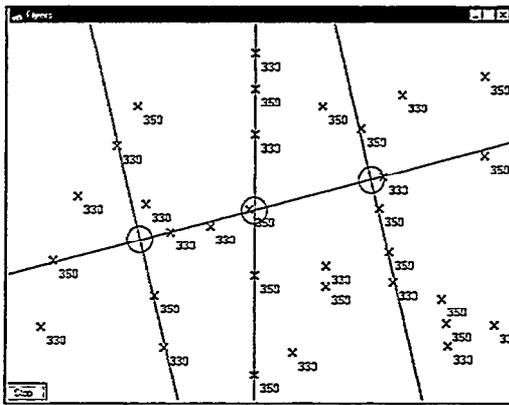


Figure 1

**Game Display with 16 Fixed Flyers and 16 Free Flyers**

A condition of the current experiment established four fixed routes with three intersections formed by these routes. This condition was based on the results of a previous experiment in which the number of the number of intersections formed by fixed routes and the number of objects not on fixed routes were manipulated in a 3 x 3 factorial repeated measures design. An additional condition of the current experiment was 16 aircraft were flying on the fixed routes of the display throughout each game.

*Procedure*

The experiment was divided into four phases: a welcome phase, an observation and training phase, a game phase, and a debriefing phase. During the welcome phase, participants completed a consent form, which provided a high-level description of the experiment, its duration, and its research benefits, and then a demographic survey was given to each participant. At the conclusion of this phase, a five minute resting baseline ECG was recorded for each participant.

At the outset of the observation and training phase, experiment instructions were read to each individual. These instructions provided a description of the object of the game, as well as a brief tutorial on how to change the altitudes of the aircraft targets. Participants were then allowed to observe a sample game for approximately 2 minutes. During this observation period, an explanation of the nature of the movement of the targets was provided so as to aid participants in reducing the number of false detections during the course of the upcoming games.

During the game phase, each participant played 5 games. Based on techniques used in previous studies, games alternated in relative difficulty between easy and hard [1].

The fixed route/Free Flyer configuration for each game is shown in Table 1.

	Game 1	Game 2	Game 3	Game 4	Game 5
Fixed Route Aircraft	16	16	16	16	16
Free Flight Aircraft	4	12	0	16	8

**Table 1**  
**The Fixed Route/Free Flyer Configuration**

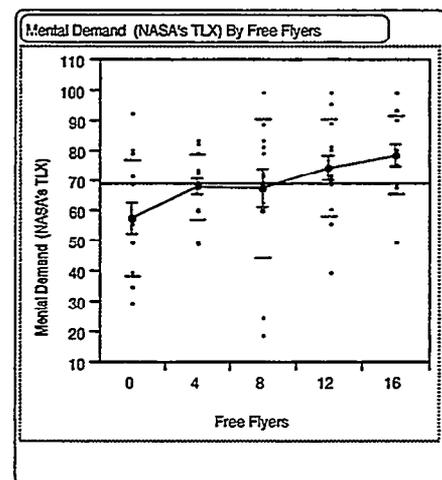
Performance statistics were compiled for each participant on the number and time of each aircraft collision per game, as well as the number and time of altitude changes. At the completion of each game, participants completed a subjective workload assessment based on NASA's Task Load Index (TLX) multi-dimensional rating procedure [5]. Participants provided responses to scales which ranged from very low to very high in six different workload areas: mental demand, physical demand, time pressure, performance, effort, and frustration.

At the conclusion of the experiment, participants were debriefed on their impressions of the experiment.

**RESULTS**

*Subjective Measure of Mental Effort*

NASA's TLX sub-scale for Mental Demand was used to validate participants' perceived mental effort across trials. A repeated measures Analysis of Variance (ANOVA) on the responses of all participants to the NASA TLX Mental Demand sub-scale demonstrated that free flyers had a significant effect on the subjective measure of mental effort ( $F(4,48) = 5.92, p=0.001$ ) (see Figure 2).



**Figure 2**  
**Mental Demand Sub-Scale Scores on NASA's TLX for All Participants (by Number of Free Flyers).**

Subjective responses on the Mental Demand sub-scale for each of the free flyer treatments were tested against the control (no free flyer treatment) for significant differences using a *t* test for non-independent samples. Contrasts between the control and the 4, 8, 12 and 16 free flyer treatments were significant for the 12 Free Flyer treatment using Dunnett's test ( $tdl=2.99$ ).

It should also be noted that the complexity factor (free flyers) had a significant effect on the subjective measure of Overall Workload,  $F(4, 48) = 7.33, p < 0.001$ . (The NASA TLX Overall Workload score is a composite of six sub-scale scores, of which Mental Effort is one).

#### Physiological Measure of Mental Effort

Results of HRV for the participants were obtained across all trials. While there was a marginally significant effect of HRV across trials for all participants ( $F(4, 48) = 2.49, p = 0.05$ ), the Group with ATC experience showed significant sensitivity to the manipulation of the independent variable,  $F(4, 16) = 4.75, p = 0.01$ . Figure 3 shows a consistent decrease in HRV for those with ATC experience as free flyers were increased from 0 to 12 (an indication of an increase in mental effort). The increase from 12 to 16 free flyers, however, resulted in an increase in HRV. An ANOVA comparison between the 12 free flyer and 16 free flyer trials yielded a marginally significant result,  $F(1, 4) = 4.21, p = 0.11$ .

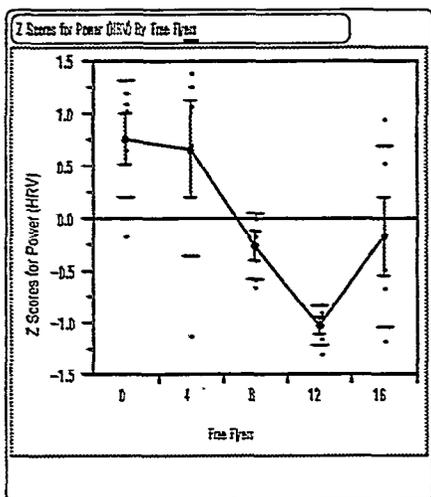


Figure 3

HRV Scores for Participants with ATC Experience (by Number of Free Flyers).

#### DISCUSSION

While the results of this study are preliminary and warrant further experimental investigation, there are two findings that indicate the potential value of HRV. The first is that HRV showed sensitivity to the manipulation of the independent variable. For the group with ATC

experience, there was significant sensitivity; for all participants there was marginal sensitivity.

A second finding is that while HRV of the ATC experienced population appeared to decrease consistently with increases in the number of free flyers in the environment (an indication of an increase in mental effort), the increase from 12 to 16 free flyers resulted in an increase in HRV. This points to the potential use of HRV as an indicator of the point at which the capacity of human resources to process the increased number of free flying targets is being exceeded. It has been found that when tasks become too difficult there is a tendency for humans to disengage from the task, resulting in an increase in HRV [1]. Thus, the increase in HRV between the 12 and 16 free flyer condition may provide an indication that the user is "losing the mental picture" of the visual display — a condition which suggests increased safety risk for the free flying and surrounding aircraft.

While not conclusive, there is evidence that suggests that the increase from 12 to 16 free flyers marks the point at which users move from what Norman and Bobrow refer to as a resource-limited condition to a data-limited condition [14]. As free flyers are increased from 0 to 12, users appear to be resource-limited as evidenced by the increase in mental effort. That is, users attempt to maintain acceptable levels of performance by increasing the amount of effort expended in keeping the aircraft targets from colliding. This increase in mental effort is supported by a corresponding decrease in HRV. As free flyers are increased from 12 to 16, users are presented with more data than they can handle with acceptable performance. This is what Norman and Bobrow call a data-limited condition; that is, more effort expended will not necessarily result in improved performance [14]. In fact, the overall increase in HRV provides an indication of a decrease of expenditure of effort and a willingness to settle for lesser performance. This phenomenon may explain the marginally significant effect for HRV obtained for all 13 participants. In other words, for those individuals not previously experienced in performing ATC tasks, even the tasks with 12 or fewer free flyers in this experiment were too difficult to be able to use HRV as a reliable discriminator for the expenditure of mental effort.

Another possible explanation for the increase in HRV between the 12 and 16 free flyer trials for the ATC experienced group is that the 16 free flyer trial did not require as much workload or mental demand as the 12 free flyer trial. NASA's TLX Overall Workload ratings and Mental Demand sub-scales scores did not support this explanation, however. Both showed an increase in the perceived workload and mental demand as free flyers were increased from 12 to 16.

### FUTURE RESEARCH

We believe that if the level of mental effort exerted by a user while interacting with an interface is reflected in HRV, then exploratory research in using HRV for interface design and analysis, as well as a component of an adaptive interface design would be logical follow-ons to the current study.

Used in conjunction with other design tools, HRV may be an aid in both forecasting and diagnosing problematic interfaces. Perhaps HRV may even be discriminating enough to identify the types of individuals who would be most troubled by certain aspects of an interface. Given the "unique" signature of an individual's HRV, the tailoring of an interface to meet the specific needs of individual users may be possible.

Assuming that HRV can be established as a reliable indicator of user state, it may also provide a research basis for its use as a component of an adaptive interface. Within the framework of current interface design methodologies, it is often known that some designs are more effective than others. However, the reason for their effectiveness may not be known. The inclusion of a physiological measure such as HRV as a part of an adaptive interface design methodology may provide useful additional insight as to what will and what will not work in the context of adaptive interface development.

Results of this study suggest conducting a more carefully designed study in which HRV is used as the basis for altering an interface. Overall task performance would be measured; correlations between HRV and performance, and between HRV and a complexity measure would also be examined. Profiles for each participant would be developed in which HRV is plotted against performance for varying measures of complexity to determine the presence of main effects and interactions. Also important, will be to ensure that the experiment design controls for order effects and other threats to internal and external validity. Finally, ensuring a sufficient number of participants is needed to increase the power of statistical tests of significance.

### ACKNOWLEDGMENTS

We thank the MITRE Corporation for providing the funding for the ECG and data analysis software used in the study. Thanks are due also to the members of the HCI group at the George Washington University for their helpful suggestions on drafts of this document. Finally, we would like to thank the referees for their valuable insights and comments that were incorporated in the final version of this paper.

### REFERENCES

1. Aasman, J., Mulder, G., and Mulder L.J.M (1987). Operator effort and the measurement of heart-rate variability. *Human Factors*, 29, 161-170.
2. Coles, M.G.H., and Sirevaag, E. (1987). Heart rate and sinus arrhythmia. In Gale, A. and Christie, B. (Eds), *Psychophysiology and the Electronic Workplace*, Chichester, UK: Wiley.
3. *Dataq Instruments WinDaq/200 User's Manual*, software release level 1, Akron, Ohio (1996).
4. Federal Aviation Administration (1995). *FAA Aviation Forecasts -- Fiscal Years 1995-2006* (FAA-APO-95-1). Washington, DC: Federal Aviation Administration.
5. Hart, S.G., and Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-178). North-Holland: Elsevier Science Publishers B.V.
6. Heart Rate Variability: Standards of Measurement, Physiological Interpretation, and Clinical Use (1996). Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation*, 93, No. 5, 1043-1065.
7. Hopman, J. C. W., Kollee, L. A. A., Stoelinga, G. B. A., van Geijn, H. P., and van Ravenswaaij-Arts, C. M. A. (1993). Heart Rate Variability. *Annals of Internal Medicine*, 118, 436-447.
8. Kramer, A. F. (1991). Physiological metrics of mental workload: A review of recent progress. In D. L. Damos (Ed.), *Multiple-Task performance* (pp. 279-328). London: Taylor and Francis.
9. Mogford, R.H., Murphy, E.D., Raske-Hofstrand, R.J., Yastrop, G., Guttman, J. A. (1993). The use of direct and indirect techniques to study airspace complexity in air traffic control. In *Proceedings of the First Mid-Atlantic Human Factors Conference* (pp. 196-202). Virginia Beach, VA.: Human Factors Society.
10. Mogford, R.H., Murphy, E.D., Raske-Hofstrand, R.J., Yastrop, G., Guttman, J. A. (1994). *Research techniques for documenting cognitive processes in Air Traffic Control: Sector complexity and decision making* (Report No. DOT/FAA/CT-TN94/3). Atlantic City International Airport, N.J.: FAA Technical Center.
11. Mogford, R.H., Guttman, J. A., Morrow, S. L., and Kopardekar, P. (1995). *The complexity construct in air traffic control* (Report No. DOT/FAA/CT-TN95/22). Atlantic City International Airport, N.J.: FAA Technical Center.
12. Mulder, L. J. M., Van Roon, A. M., and Schweizer, D. A.. *Cardiovascular Data Analysis Environment*

- (CARSPAN) *User's Manual*, Groeningen, The Netherlands (1995).
13. Mulder, L. J. M., and Mulder, G. (1987). Cardiovascular reactivity and mental workload. In O. Rompelman and R. I. Kitney (Eds.), *The beat-by-beat investigation of cardiovascular function* (pp. 216-253). Oxford, UK: Oxford University Press.
  14. Norman and Bobrow (1975). Data and resource processing limits. *Cognitive Psychology* (7), 44-64.
  15. RTCA (1995, October). Final Report of RTCA Task Force 3 *Free Flight Implementation*. Washington, D.C.: RTCA Task Force 3.
  16. Sirevaag, E. J., Kramer, A. F., Wickens, C. D., Reisweber, M., Strayer, D. L., and Grenell, J. F. (1993). Assessment of pilot performance and mental workload in rotary wing aircraft. *Ergonomics*, 36, 1121-1140.
  17. Tattersall, A., and Hockey, G. (1995). Level of Operator Control and Changes in Heart Rate Variability during Simulated Flight Maintenance. *Human Factors*, 1995, 37(4), 682-698.
  18. Tulga, M.K., and Sheridan, T.B. (1980). Dynamic decisions and workload in multitask supervisory control. *IEEE Transactions on Systems, Man, and Cybernetics*, 1980, 10, 217-232.
  19. Velichkovsky, B. M., and Hansen, J. P. (1996). New technological windows into mind: There is more in eyes and brains for human-computer interaction. In *CHI 96 Proceedings* (pp. 496-503). Vancouver, BC Canada.
  20. Vicente, K.J., Thornton, D. C., and Moray, N. (1987). Spectral Analysis of Sinus Arrhythmia: A Measure of Mental Effort. *Human Factors*, 1987, 29(2), 171-182.
  21. Wierwille, W. W., and Eggemeier, F. T. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors*, 35(2), 263-281.
  22. Wilson, G. F. (1993). Air-to-Ground training missions: A psychophysiological workload analysis. *Ergonomics*, 36, 1071-1087.
  23. Wilson, G.F., and Eggemeier, F. T. (1991). Psychophysiological assessment of workload in multi-task environments. In D. L. Damos (Ed.), *Multiple-task-performance* (pp. 329-360). London: Taylor and Francis.