PERIODICA POLYTECHNICA SER. SOC. MAN. SCI. VOL. 7, NO. 2, PP. 119–134 (1999)

VALIDATING A NEW METHOD FOR ERGONOMIC EVALUATION OF HUMAN COMPUTER INTERFACES

Lajos IZSÓ[1], Gábor MISCHINGER and Eszter LÁNG

Department of Ergonomics and Psychology
Budapest University of Technology and Economics
H–1521 Budapest, Hungary

Received: March 11, 2000

Abstract

Human Computer Interaction (HCI) is the basic component of using all kinds of information technological systems. This is why examination methods capable of assessing factors influencing users' performance, current mental effort and well-being during HCI in a well established and scientifically sound way have got great importance. The 'current' attribute for mental effort means here that in addition to overall measures concerning a relatively longer period – several hours or a whole working day – such methods should also provide informative and valid data about users' mental investments at the temporal resolution corresponding to users' elementary actions as keystrokes, mouse clicks, etc. It has already been shown that under certain circumstances Heart Period Variability (HPV) could be a measure of momentary mental effort. This paper gives a short overview of applications of HPV in ergonomics in general, and based on empirical evidence intends to prove that this methodology, after a careful adaptation, could be an especially adequate and powerful technique for monitoring mental effort in HCI. The paper outlines the main principles of a new method and the related components of the integrated system (INTERFACE) developed by us for investigating HCI from several aspects with emphasis on assessing mental effort. A detailed application example is also provided.

Keywords: Human Computer Interaction, mental effort, user performance, user behaviour, Computerized Directory Assistance Services.

1. Introduction

Components of Human Computer Interaction (HCI) are the basic acts of using information technological systems, all these systems are used by certain kinds of users, who somehow operate them and get into contact with them. Nowadays it is not an exaggeration to talk about 'information revolution' as information technologies not only radically change the world of labour, but also quickly penetrate into the domain of home activities, leisure time and hobby, telecommunication, sports, education, transport, etc. An ergonomically well designed User Interface (UI) can make the work of thousands easier, while a bad design could cause less or greater annoyance, embitter whole working days or even can result in significant losses or damages for the same number of users. As a result of the foregoing an urging

[1]The first author acts as designated contact person, Mail: Egry J. u. 1. Bldg. E III/11, Budapest, Hungary–1111, Telephone: 36 1 463–2653, Fax: 36 1 463–2106, E-mail: izsolajos@erg.bme.hu

need has been articulated for methods capable of monitoring users' characteristic features, especially their current mental effort during the HCI validly and reliably in an '*on line*' way.

The reason is that in possession of such methods it would be possible to identify the 'weak points' of the actual software packages or other smart products – hopefully still in prototype stage – and they could be modified or redesigned based on well established principles.

As in HCI the 'event flux' is generally very high compared with more traditional activities, that is during a relatively short period of time quite a lot things could happen between the human and the computer, the methods based on users' reports have their own specific limits. Even after a short time users simply cannot recall the great number of details of usability or navigation problems they encountered. Furthermore, having no basis for reference, they can take certain difficulties as inevitable or to be attributed to their own inexperience. This is, of course, not to say that the methods based on users' reports are useless in the research of HCI. On the contrary, they are indispensable in their place: users' global impressions or remarks on some conspicuous details usually bear important and useful information.

Thus, there is a definite need for methods capable of providing informative and valid data about users' mental investments at the fine temporal resolution corresponding to users' elementary actions as keystrokes, mouse clicks, draggings, noticing new screen messages, etc. Based on empirical evidence in this paper Heart Period Variability (HPV) is proposed to be used as a measure of computer users' momentary mental effort.

2. INTERFACE: a Complex Method for Testing Efficiency of Human–Computer Interaction

2.1. Fundamentals of the Method

As it has been stated elsewhere (IZSÓ and ZIJLSTRA, 1997) in working activity efficiency refers to the ratio between 'benefits' and 'costs':

$$\text{Efficiency} = \frac{\text{Benefits}}{\text{Costs}}.$$

If 'benefits' and human 'costs' are properly operationalised for the case of human–computer interaction by using appropriate measures adequate both to the particular task of interest and the people working on it, the '*Efficiency*' concept – as has been proved – can be a powerful indicator. The potential strength of this approach lies in the fact that it simultaneously takes into consideration both sides of the same coin.

A complex method has been developed to identify those parts of human–computer dialogues that require the most mental effort – the higher human 'costs' – from users. This method simultaneously investigates:

a) the performance of users (based on speed, performance time, number of keystrokes hit and other performance measures)
b) the behaviour and observable actions of users while performing (based on the following two pictorial information: (1) video recording of users' behaviour (acts, mimics, gestures, etc.) during work with the given software interface, and also (2) video recording of the current screen content)
c) a special psycho-physiological parameter – the power spectrum of HPV – which is associated with mental effort, and applied as an objective measure of current mental effort.

In terms of the efficiency concept mentioned above point a) refers to 'benefits', while some measures of the 'costs' can be derived from points b) and c).

In addition to observable elements of behaviour the proposed complex method comprises also the use of questionnaires and interviews assessing mental models, subjective feelings and opinions of users about their perceived task difficulty and fatigue.

Measuring and data collecting facilities are integrated into the user interface testing workstation called INTERFACE[2] the conceptual arrangement of which can be seen in *Fig. 1*. The basic advantage and novelty of this methodology lies in

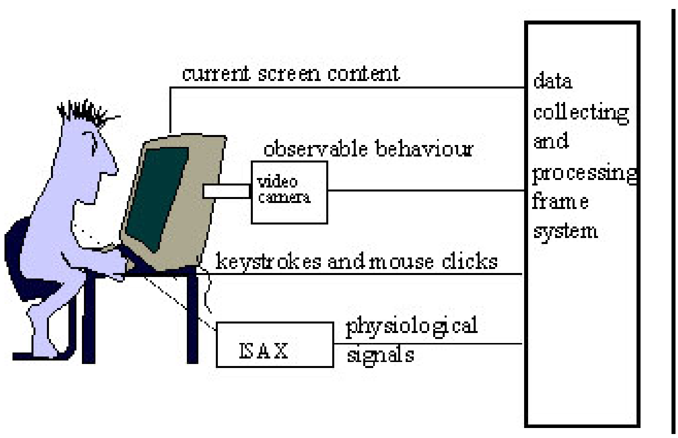

Fig. 1. Conceptual arrangement of the INTERFACE user interface testing workstation

its capability of recording continuous on-line data characterising the user's current mental effort. In this paper therefore we only list the basic ideas of assessing users' performance and behaviour, and later on a more detailed description of HPV analysis is given and its relation to current mental effort actually invested by the user is shown in more details. In section 3 an example application is presented with supporting evidence concerning the validity of the method.

[2] INTegrated Evaluation and Research Facilities for Assessing Computer-users' Efficiency.

2.2. Assessing Performance and Behaviour of Users

Performance measures have to be defined by taking into account particular characteristics of both the software product to be tested and the preparedness of actual users. These measures are generally numbers of countable things performed during a period of time such as clients served, forms filled in, characters typed in, items entered into a register or monthly returns, etc. Performance criteria often closely relate to time, in these cases certain predefined amounts of work must be done during a unit of time. While measures mentioned above could be considered as measures of quantity, performance measures based on the errors committed could be taken as measures of quality of computer work.

Recording *users' behaviour* has outstanding importance for the following reasons. First of all, by recording only performance measures we cannot get any information about difficulties users encountered. In this respect video-recording of the user's face and activity is an extremely rich source of psychological information as it directly reflects the mental state (e.g. boredom, routine activity in familiar environment, attention-demanding task, helplessness, getting lost, emotions as frustration, joy, boredom, etc.). Together with performance measures it makes possible identifying relatively weak points of the particular software interface.

Video-recorded behaviour has also been proven to be an indispensable tool for identifying the possible artefacts during HPV analysis. Namely, HPV is sensitive to large muscle movements and changes in breathing frequency and amplitude as well. Although computer work is typically done in seated position and therefore generally no large muscle movements occur, sometimes people may turn around, stretch themselves out or even stand up for a minute. Similarly they can talk, yawn, laugh or sneeze which may result in radical changes in their breathing frequency. Finally, video-recordings have made possible studying important pieces in more details by playing back events as often as it was necessary.

2.3. Assessing Mental Effort Via Analysing Users' HPV Power Spectrum

2.3.1. An Overview of Applications of HPV in Ergonomics

A number of studies has shown that increasing mental load causes a decrease in heart rate variance (KALSBEEK and ETTEMA, 1963; LUCZAK and LAURIG, 1973; MULDER and van der MEULEN, 1973; ROHMERT et al., 1973; SEKIGUCHI et al., 1979; MULDER, 1980; MULDER and MULDER, 1981; LEE and PARK, 1990; LÁNG and SZILÁGYI, 1992; LÁNG et al., 1994; ZIJLSTRA, 1993; WIETHOFF, 1997; IZSÓ and WIETHOFF, 1997).

Cardiac interbeat interval sequences show spontaneous quasi-periodic fluctuations. This heart rate (HR) variability is the result of the superimposition of different sources of variation. The variation of HR is mainly mediated by the sinus node which is the pacemaker of the heart. It is innervated by the sympathetic nerves to the heart and by the parasympathetic *nervus vagus*. So the HR variation is largely determined by a balance between levels of activity in the cardiac sympathetic and parasympathetic nerves. Analysis of HR variance in the frequency domain allows separate contributions to be isolated.

Depending on recording techniques, methods in signal-processing and spectral analysis some authors prefer to use variability of instantaneous heart rate (beat/minute - bpm), while others variability of inter-beat intervals, that is, heart periods (msec). Physiological conclusions, however, drawn from spectral analysis of heart rate variability (HRV) or heart period variability (HPV) can be used interchangeably.

It has been shown (HYNDMAN et al., 1971; LUCZAK and LAURIG, 1973; MULDER and van der MEULEN, 1973; SAYERS, 1971, 1973; WOMACK, 1971; AKSELROD et al., 1981, 1985; PAGANI et al., 1986; WEISE et al., 1987) that short-term fluctuations (in the time scale of seconds to minutes) in heart period (HP) were concentrated in three principal peaks of the heart period variance (HPV) power spectrum.

The high-frequency (HF) peak of the HPV power spectrum is centred around the respiratory frequency of the species studied. In humans it typically appears between 0.15 and 0.45 Hz. It reflects the so called respiratory sinus arhythmia (RSA), a phenomenon described by Ludwig in the last century: during inspiration the heart rate increases, while during expiration it decreases. Experimental and clinical investigations demonstrated that RSA is vagally mediated. Parasympathetic blockade essentially abolishes respiratory frequency fluctuations in HR while beta or alpha sympathetic blockade did not affect the HF component (AKSELROD et al., 1985; MCCABE et al., 1981; PORGES et al., 1981; POMERANZ et al., 1985).

The mid-frequency (MF) peak centred around 0.1 Hz reflects vasomotor blood pressure fluctuations. HR fluctuations in the MF range may reflect a baroreceptor response to these blood pressure fluctuations in this frequency band (HYNDMAN et al., 1971, WESSELING and SETTELS, 1985). Experiments involving pharmacological manipulations and postural changes show that contrary to the HF (respiratory) peak mediated solely by the parasympathetic system, oscillations at lower frequencies are mediated jointly by the parasympathetic and sympathetic nerves (AKSELROD, 1985, 1988; WEISE, 1987; POMERANZ, 1985; PAGANI, 1986; SAUL et al., 1990).

SAYERS (1971, 1973) found that during mental workload consistent changes occur in the heart period (HP) spectrum especially in the band from 50 to 150 mHz. According to MULDER et al. (1973) the mid-frequency (MF) band of HPV (70–140 mHz) appeared to be more sensitive to mental workload than the total variance or respiratory fluctuations. Suppression of the MF-component of HPV spectra reflects the effort requested by the mental task. The rebound of the spectral power during relaxed period following the completion of the mental task is related to the degree of preceding effort. The greater the rebound the more pronounced the preceding effort. It is believed that mental load (when the task requires explicit effort) operates like a defence reaction. The defence reaction is characterised by a decrease in sensitivity of the baroreflex which results in a decrease of heart rate variance (HRV) because changes in the blood pressure will be less reflected in changes in heart rate. Defence reaction involves suppression of the vagal component of the reflex (see MULDER, 1980).

For more about spectral analysis of HPV see the review by LÁNG and SZILÁGYI (1991).

2.3.2. *An Instrument (ISAX) for Assessing Power Spectra of HPV*

To assess spectral components of HPV power spectra in field studies an integrated system has been developed called ISAX[3] (LÁNG and HORVÁTH, 1994, LÁNG et al., 1994a, LÁNG et al., 1997, 1998). It consists of a portable equipment capable of 24 hour measurement and storage of heart periods (that is: *R* to *R* intervals of the electrocardiogram) beat by beat, respiratory and optionally other bio-signals, and a software package for spectral analysis of the data stored in a single IBM PC and algorithms to evaluate parameters of the significant spectral components of the power spectra of HPV, and also plain text table output for further statistical processing as well.

The acquisition module is a small (300 g) portable plastic box that can be mounted on the patient by a clip, and connected to the sensors. The ambulatory recorded data are stored in the built-in NVRAM. The recorded data are read and processed by a host computer. For electrocardiogram recording precordial bipolar leads are used. Three electrodes (exploring, indifferent, ground) have to be placed on the chest. The exploring (positive) electrode on the 5^{th} rib or in the fifth intercostal space on the left medioclavicular line (V5-position), the indifferent (negative) electrode high up on the sternum, and the ground electrode in the sixth intercostal space on the left median auxiliary line. Other ECG electrode locations can also be selected, if desired, so as to maximise *R*-wave magnitude.

ISAX provides 2 channels of respiration. The mechanical motion of the chest, thus, respiration amplitude and frequency can be measured with the strain gauge method (solid state resistive strain gauge which responds linearly to changes in abdominal or thoracic circumference). The strain gauge belt has to be attached around the patient's lower rib-cage. Nasal airflow can also be measured with the nasal thermistor airflow sensor under the nostrils.

Two channels serve the purposes of event-marking in order to be able to identify data sequences recorded in special conditions. There is one push button mounted on the module's front panel, and there exists a possibility for external trigger input allowing remote marking.

R waves of the recorded ECG are detected using a special software, then *R* to *R* intervals are measured and stored. The data transfer between the module and the PC can be carried out by using one of the standard PC's peripheries (serial COM1, or COM2 ports) by the reader software producing structured data for immediate analysis.

[3] Integrated System for Ambulatory Cardio-respiratory data acquisition and Spectral analysis

Processing of RR-interval series by the ISAX program consists of steps as follows. First RR-interval series are interpolated for the sampling procedure (1 Hz). RR-interval time functions are displayed for interactive selection of appropriate analysis frame. A sufficiently long stationary and representative part of the RR-interval function is selected for spectral analysis. The RR-interval function marked for analysis is converted into zero-mean process. An all-pole Auto-Regressive model is fitted to the data set using a modified Burg algorithm (GRAY and WONG, 1980). Model order is selected (between 10 and 12) which minimised the Akaike's final prediction error (AKAIKE, 1969). Central frequencies and bandwidths of the spectral peaks are computed from the numerically determined poles of the synthesis model. Spectral decomposition is undertaken, and spectral components are assigned to one of the three bands: low-frequency (LF) (10 to 70 mHz), mid-frequency (MF) (70 to 150 mHz), high-frequency (HF) (150 to 450 mHz). The following non-spectral and spectral parameters are computed for further statistical analysis: *RR-interval mean* (ms) (mean of the analysis frame), the corresponding standard deviation – *RR-interval SD* (ms) (standard deviation of the analysis frame), *LF-power* (ms^2), *MF-power* (ms^2), and *HF-power* (ms^2) of HPV spectrum, and corresponding central frequencies (CF) (mHz).

2.2.3 Characterisation of the 'Profile'

One of the problems of spectral analysis is the issue of non-stationarity of the time series to be analysed. Interpretation of the outcomes is hindered by dependency on the chosen segmentation of the time series to be analysed. In the program package of ISAX this problem has been attacked by an approach called by MULDER et al., (1988) 'spectral profile' method. Spectral computations are carried out on short time segments such as 30 or 60 seconds. By shifting such segments over the time series to be analysed, and by introducing a certain overlap (e.g. 50 up to 90% overlap) a series of spectral values is obtained indicating a varying spectral power in a selected frequency band. The segment length and the overlap factor can be adjusted in a wide range. This way spectral power of MF component of HPV versus time is obtained. A useful feature of the software package of ISAX is simultaneous visualisation of the time function of RR-intervals in the upper window with the same time scale. Thus changes in heart rate can be tracked together with simultaneous changes in spectral estimates of HPV. This may help in giving deeper insight into the autonomic nervous mechanisms of subjects during human–computer interaction.

3. An Example Application of the Method

3.1. The User Interface of the Software to be Tested

The Hungarian Telecommunication Company (MATÁV Rt.) launched a big project in 1996 aiming at establishing a centralised national inquiry system titled Computerized Directory Assistance Services (CDAS). The software component of this system was developed and adapted by IBM Hungary. As the company was interested in identifying relatively weak points of the user interface in the areas of dialogue flexibility, usability of search functions, validity of data base, lucidity of hit lists, etc. they gave an order for a thorough ergonomic evaluation of the CDAS software. Authors of this paper took part in this real work field study.

3.2. Hypotheses

1. By the help of the INTERFACE method it is possible to identify the relative weak points of the system.
2. Identified weak points are associated with operators' higher observable mental effort.
3. Operators' HPV would be more depressed while answering calls demanding higher mental effort than while answering not demanding calls.

3.3. Subjects

Subjects of the study were 15 CDAS operators who were performing their normal job. Of the operators 14 were female and only 1 male, which represented the real female/male ratio in this job. Their average age was 27, with the age distribution of 20–25 60%, 26–30 13%, above 30 27%. As an average they were employed by the company for 7.5 years, in their present job as inquiry service operators for 4.5 years. At the time of the study they had been using the CDAS software only for 1–2 months. The operators were informed that certain data were going to be recorded on them and were asked to allow placing ECG electrodes on their chest for recording by the data collecting module of the ISAX.

3.4. Procedure

Each operator's interaction with the CDAS software was recorded for a period of 50 minutes, half of which was with "real" incoming calls from "real" clients. In addition to this real life situation with calls of random demands during the other 25 minutes the operators received 43 carefully chosen 'fictive' calls from our helpers. After the 'fictive' session the operators were asked to judge each call, that is task situation, as difficult, medium or easy. Some of these "fictive" calls were judged by the majority of operators as difficult, while others as easy. There were, of course, certain individual differences: some calls were judged differently by different operators.

The following time data were automatically recorded and later on analysed:

TOTAL: total time (from the first moment of incoming call to completing answering client fully satisfying him or her),

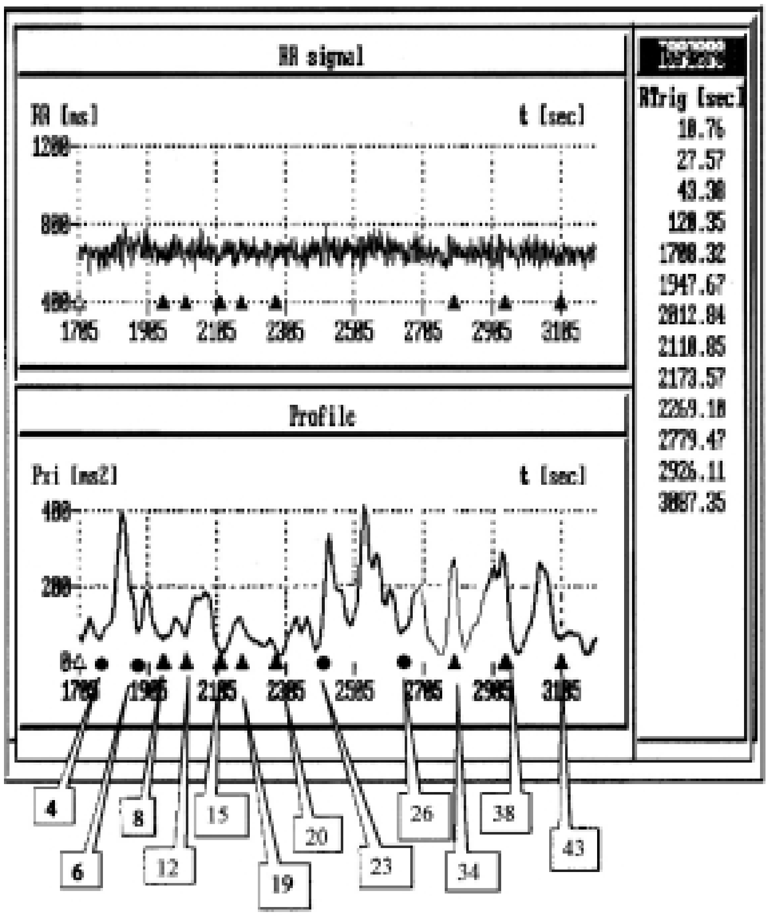

Fig. 2. A typical example profile of an operator. The upper curve is the RR interval – time function, the lower is the MFP as a function of time. The numbers along the time axis – scaled in seconds – are corresponding to the situations judged as *difficult*. As can be seen this operator had 12 *difficult* calls, including the 7 ones belonging to the selected set.

QUESTION: question time (from the first moment of incoming call to operator's first keystroke),

SEEK seeking time (from operator's first keystroke to starting seek function),

FIND finding time (from starting seek function to appearing first hit list),

ANSW answering time (from appearing first hit list through communicating it to acknowledging it by the – hopefully satisfied – client).

TOTAL = QUESTION + SEEK + FIND + ANSW

For testing hypothesis 1) these time data and the possible ways of their reductions were analysed. The regularly slow steps of the whole event chain were considered as identified weak points (‘*bottlenecks*’). To test if these were associated with operators’ higher observable mental effort – as stated by hypothesis 2) – direct observation, video recording and interviews supported by video-playback were used. In order to test the hypothesis 3) that largest set of operators was identified, who had the same opinion concerning a relatively large number of calls. So finally 9 operators were identified out of the 15, whose judgements were all ‘difficult’ concerning the calls 4, 12, 15, 19, 34, 38, and 43 and at the same time whose judgements were all ‘easy’ concerning the tasks 18, 24, 30, 31, 39, and 42. The other calls were judged either as “medium”, or differently.

This way 7 *difficult* and 6 *easy* calls were identified for 9 operators. These categories were also confirmed by our observation.

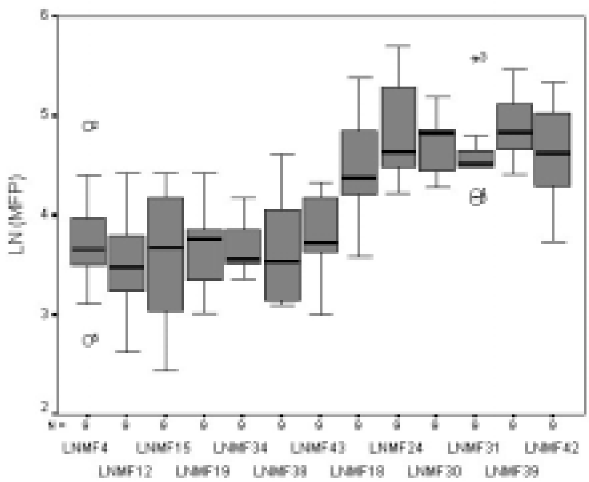

Fig. 3. The distributions of the 9 transformed data for each *difficult* and 6 *easy* calls in a box-plot. Note that all the values of Ln(MFE) of *difficult* calls (D) are much lower than those of *easy* calls (E).

3.5. Research Instruments

The main hardware and software components of the INTERFACE (*Fig. 1*) were installed in the huge operator room of the CDAS and were used in the study. In addition to that some pieces of the infrastructure of the CDAS working place also were used as local computers, ‘artificial’ telephone lines, etc. For our present purposes the most important component of the INTERFACE was the ISAX system described in details in section 2.3.2. Using the ISAX required placing ECG electrodes to the operators' chest and attaching the data collecting module. This module recorded the heart periods (intervals between two successive *R* waves of the ECG). Spectral components of HPV were calculated by the help of ISAX software.

Table 1. Mean time components of the event chain

TIME COMPONENTS	PERCENTAGES	MEAN (s)
QUESTION	8.7%	2.45
SEEK	36.9%	10.40
FIND	27.6%	7.77
ANSWER	26.8%	7.55
TOTAL	100	28.17

3.6. Results

Mean time components of the whole event chain are shown in *Table 1* in percentages. It can be seen that a better support for data entry – SEEK – could speed up the process. Similarly, a better arranged – more clear-cut – hit list could reduce FIND.

A more detailed analysis identified some particular ways, in which – taking into account the interaction between the operator, the data basis. the hit list, and hardware components – it was really possible to shorten the steps SEEK and FIND.

Such a way, as stated by hypothesis 1), by the help of the INTERFACE method it was possible to identify relative weak points of the system.

Table 2. Results of the ANOVA

Tests of Within-Subjects Effects

Measure: MEASURE_1
Sphericity Assumed

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
DIFFIC	32.280	12	2.690	15.835	0.000	190.016	1.000
Error(DIFFIC)	16.309	96	0.170				

[a] Computed using alpha = 0.05

It was also proved by observations and interviews, that in accordance with hypothesis 2), the identified weak points in the majority of cases were associated with operators' higher observable mental effort. In practice these situations were the following: operators were disturbed and embarrassed by not up-to-date, overlapping

Table 3. Results of the Contrast Analysis

Test of Within-Subjects Contrasts

Measure: MEASURE_1

Source	Transformed Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[a]
DIFFIC	DIFFIC_1	1.468	1	1.468	6.715	0.032	6.715	0.623
	DIFFIC_2	3.594	1	3.594	20.146	0.002	20.146	0.974
	DIFFIC_3	2.882	1	2.882	13.584	0.006	13.584	0.896
	DIFFIC_4	0.974	1	0.974	3.518	0.098	3.518	0.379
	DIFFIC_5	1.736	1	1.736	17.689	0.003	17.689	0.956
	DIFFIC_6	4.923	1	4.923	51.338	0.000	51.338	1.000
	DIFFIC_7	2.793	1	2.793	37.361	0.000	37.361	1.000
	DIFFIC_8	2.034	1	2.034	12.417	0.008	12.417	0.869
	DIFFIC_9	1.744	1	1.744	30.414	0.001	30.414	0.998
	DIFFIC_10	1.951	1	1.951	12.486	0.008	12.486	0.870
	DIFFIC_11	5.078	1	5.078	44.642	0.000	44.642	1.000
	DIFFIC_12	1.902	1	1.902	7.614	0.025	7.614	0.677
Error(DIFFIC)	DIFFIC_1	1.749	8	0.219				
	DIFFIC_2	1.427	8	0.178				
	DIFFIC_3	1.697	8	0.212				
	DIFFIC_4	2.216	8	0.277				
	DIFFIC_5	0.785	8	0.812E-02				
	DIFFIC_6	0.767	8	0.590E-02				
	DIFFIC_7	0.598	8	0.476E-02				
	DIFFIC_8	1.311	8	0.164				
	DIFFIC_9	0.459	8	0.734E-02				
	DIFFIC_10	1.250	8	0.156				
	DIFFIC_11	0.910	8	0.114				
	DIFFIC_12	1.998	8	0.250				

[a] Computed using alpha = 0.05

and clumsy data basis items, sometimes they tried to work 'by heart' – based their memories – or had to ask help from a more experienced operator. Similarly, a badly arranged hit list, even if it contained the required data, could seriously confuse the operator. As they knew that their performance was measured in terms of time, wasting time because of the poor software support just made them even more anxious.

Spectral power of the mid-frequency band of HPV (MFP[4]) versus time – the so called 'profile' mentioned in section 2.3.3. – has been examined.

The minima of MFP in the individual profiles, practically without exception, corresponded to the short time periods, where high mental efforts were both ob-

[4]Mid-Frequency Power

Table 4. A summary of interaction problems demanding high mental effort and the reported difficulties. The 'Call No' corresponds to the event numbers of the typical example profile shown in *Fig. 2*.

Call No.	Nature of problem	Reported difficulty confirmed also by observation
4.	'74/311-04..., do not know the last digit, Császár'	Does not know the *joker* search function
6.	'Budapest, Gyergyák György, first or third district'	Too many *Gyergyák*s in both districts.
8.	'Komló, Eötvös u. 12.'	Spelling: *Ötvös* or *Eötvös*?
12.	'Budapest, playing house'	Everyday names as 'playing house' are not supported by the system, they are not in data base. The solution could have been the official name 'Palace of Wonders'.
15.	'Budapest, Andrássy u. 24'	Spelling: *Andrási*, *Andrásy* or *Andrássy*?
19.	'Dombóvár, Hunyadi u. 19'	Data base does not differentiate between street, road, and square.
20.	'Miskolc, Thury László'	Spelling: *Túri, Túry, Thúry* or *Thury*?
23.	'Budapest, Ikarusz Vehicle Manufacturing Plc.'	Too many companies have the name *Ikarusz*.
26.	'Lady tailor Dombóvár'	Profession 'tailor' and family name 'Tailor' washed together. Result: all Tailors and tailors of the town listed.
34.	'Kindergarten in Mecsekjános'	Everyday names are not supported by the system, they are not in data base. The solution could have been the official name of this kindergarten.
38.	'Ópusztaszer, Feszty panorama picture'	Everyday names are not supported by the system, they are not in data base. The solution coulds have been the official name 'Historical Memory Park'.
43.	'My car broke down in Budapest 14. district. What is the number of a car mechanic'?	System's message: 'too long list'. There are many car mechanics in the 14. District, but no one's number was provided.

served by us and reported by the subjects during the interviews after the sessions supported by video playback.

A typical example profile of an operator can be seen in *Fig. 2*.

In order to get a statistical proof for this clearly observable correspondence a repeated measure ANOVA (with 'DIFFIC' = difficulty as within subject factor) was carried out on the MFP data recorded during the 7 *difficult* and 6 *easy* calls for the 9 operators. As the distribution of the MFP data was far from being normal, first a logarithmic transformation was carried out which provided the normality. The ANOVA, therefore, was applied not directly on the MFP data, but their natural logarithms instead.

The distributions of the 9 transformed variables are shown for each *difficult* and *easy* calls in a box-plot in *Fig. 3*.

Results of the ANOVA (*Table 2*) show highly significant main effects for the DIFFIC factor within subjects.

Results of the '*deviation*' Contrast Analysis carried out after the ANOVA are presented in *Table 3*. Deviation contrasts were calculated between each level of the DIFFIC factor and the grand mean: all differences were significant.

Finally *Table 4* allows to take a closer look to the nature of problems that occur during the *difficult* calls.

3.7. Conclusions

All hypotheses have been supported, the INTERFACE has proved to be a powerful tool for analysing HCI in general.

In particular, there were also interesting findings concerning MFP as mental effort monitor. On the basis of the results obtained by ANOVA and by Contrast Analysis, taking also into account that all the values of LN (MFP) for *difficult* calls are lower, and those of *easy* calls are higher than the grand mean (refer *Fig. 3*), it follows that the differences between the LN (MFP) of *difficult* and *easy* calls are even more significant. This finding is a proof for the hypothesis 3).

So it really can be stated, that during tasks demanding more mental effort (*difficult* calls) operators' spectral power in the mid-frequency band of HPV is more depressed than during tasks demanding less effort (*easy* calls).

References

[1] AKAIKE, H. (1969): Fitting Autoregressive Models for Prediction, *Ann. Inst. Statist. Math.*, Vol. 21, pp. 243–247.

[2] AKSELROD, S. (1988): Spectral Analysis of Fluctuations in Cardiovascular Parameters: a Quantitative Tool for the Investigation of Autonomic Control. TIPS, January, 9: 6-9, Elsevier Publications, Cambridge.

[3] AKSELROD, S. – GORDON, D. – MADWED, J. B. – SNIDMAN, N. C. – SHANNON, D. C. – COHEN, R. J. (1985): Hemodynamic Regulation: Investigation by Spectral Analysis. *Am. Journal of Physiology*, Vol. 249, H867-H875.

[4] AKSELROD, S. – DORDON, D. – UBEL, F. D. – SHANNON, D. C. – COHEN R. J. (1981): Power Spectrum Analysis of Heart Rate Fluctuations: A Quantitative Probe of Beat-to-beat Cardiovascular Control. *Science*, Vol. 213, pp. 220–222.

[5] ANTALOVICS, M. – BÁNHIDI, L. – DOMBI, I. – IZSÓ, L. – LÁNG, E. – MAGYAR, Z. – MAJOROS, A. – MITSÁNYI, A. – MOLNÁR, L. (1995): New Measuring of Combined Effects. *Healthy Buildings '95, Milano, Italy. Proceedings of the 4th International Conference*, Vol. 3, pp. 1425–1430, (September 10-14, 1995).

[6] BÁNHIDI, L. – ANTALOVICS, M. – IZSÓ, L. – MAGYAR, Z. – MOLNÁR, L. (1993): Some New View Points of Laboratory Measurements with Subjects. *INDOOR AIR'93, Helsinki, Finland. Proceedings*, Vol. 1, pp. 749–754, (July 4-8, 1993).

[7] GRAY, A. H. – WONG, D. Y. (1980): The Burg Algorithm for LPC Speech Analyzis Synthesis. *IEEE Tr. ASSP.*, Vol. ASSP-28, No. 6, Dec. 1980. pp. 609–615.

[8] HYNDMAN, B. W. – KITNEY, R. I. – SAYERS, B. MCA. (1971): Spontaneous Rhythms in Physiological Control Systems. *Nature*, Vol. 233, 5B18:339.

[9] IZSÓ, L. – LÁNG, E. – BÁNHIDI, L. – MAGYAR, Z. (1995): Ambulatory Assessment of Mental Effort by a Computer Supported System. *Seventh European Congress on Work and Organizational Psychology*, Győr, Hungary. Proceedings, p. A92, (19-22 April, 1995).

[10] IZSÓ, L. – ZIJLSTRA, F. (1997): Efficiency in Work: An Approach to Interface Evaluation and Design. *8th European Conference on Work and Organizational Psychology*. Proceedings pp. 7–19, April 2-5, 1997, Verona, Italy.

[11] IZSÓ, L. – ANTALOVICS, M. – VERMEEREN, A. (1992): The Role of Pilot Studies in User-interface Research. *ECCE 6. Sixth European Conference on Cognitive Ergonomics*, Balatonfüred, Hungary. Proceedings pp. 205–216, (6-11. September 1992).

[12] IZSÓ, L. – den BUURMAN, R. (1992): A User Interface Evaluation Case Study. *Scientific Conference of the Faculty of Natural and Social Sciences*, TUB, Budapest, Hungary (28-30. September 1992).

[13] IZSÓ, L. – LÁNG, E. (1998): A Hybrid Psychophysiological Model of Mental Effort. *24th International Congress of Applied Psychology, Symposium: Sociopolitical Issues in Stress Research*. Program Book p. 60, August 9-14, 1998, San Francisco, U.S.A.

[14] IZSÓ, L. – WIETHOFF, M. (1997): Some Empirical Findings on Heart Period Variability as Measure of Mental Effort in Human Computer Interaction. *13th Triennial Congress of the International Ergonomics Association*. Proceedings Volume 5, pp. 350–352, June 29 - July 4, 1997, Tampere, Finland.

[15] KALSBEEK, J. W. H. – ETTEMA, J. H. (1963): Scored Regularity of the Heart Rate Pattern and the Measurement of Perceptual or Mental Load. *Ergonomics*, Vol. 6, p. 306.

[16] LÁNG, E. – HORVÁTH, G. (1994): Integrated System for Ambulatory Cardio-respiratory Data Acquisition and Spectral Analysis (ISAX) User's Manual. Budapest.

[17] LÁNG, E. – BÁNHIDI, L. – ANTALOVICS, M. – IZSÓ, L. – MITSÁNYI, A. – ZSUFFA, A. – MAGYAR, Z. – HORVÁTH, GY. – SLEZSÁK, I. – MAJOROS, A. – DOMBI, I. – MOLNÁR, L. (1994): A Complex Psychophysiological Method to Assess Environmental Effects (-Temperature, Illumination, Sound-) on Objective and Subjective Parameters of Humans in Simulated Work Setting. *Healthy Buildings'94. Proceedings of the 3rd International Conference* (Budapest, Hungary, 22-25 August, 1994) Vol. 2, pp. 799–803. Eds: L.Bánhidi, J. Farkas, Z. Magyar, P. Rudnai.

[18] LÁNG, E. – CAMINAL, P. – HORVÁTH, G. – RJANÉ, R. – VALLVERDÚ, M. – SLEZSÁK, I. – BAYÉS DE LUNA, A. (1998): Spectral Analysis of Heart Period Variance (HPV) – a Tool to Stratify Risk Following Myocardial Infarction. *Journal of Medical Engineering and Technology*, Vol. 22 No. 6.

[19] LÁNG, E. – HORVÁTH, G. – SLEZSÁK, I. (1997): Integrated System for Ambulatory Cardio-respiratory Data Acquisition and Spectral Analysis. *World Congress on Medical Physics and Biomedical Engineering*. Nice, France, Sept 14-19 1997. Abstract: *Medical and Biological Engineering and Computing*. Vol. 35, Supplement Part I, p. 118.

[20] LÁNG, E. – SZILÁGYI, N. (1991): Significance and Assessment of Autonomic Indices in Cardiovascular Reactions. *Acta Physiologica Hungarica*. Vol. 78, (3) pp. 241–260.

[21] LÁNG, E. – SZILÁGYI, N. – MÉTNEKI, J. – WEISZ, J. (1992): Effects of Mental Load on the Spectral Components of Heart Period Variability in Twins. *Acta Biochim. Biophys. Hung.* Vol. 26: (1-4), pp. 111–120.

[22] LEE, D. H. – PARK, K. S. (1990): Multivariate Analysis of Mental and Physical Load Components in Sinus Arrhythmia Scores. *Ergonomics,* Vol. 33, No. 1, pp. 35–47.
[23] LUCZAK, H. – LAURIGM, W. (1973): Analysis of Heart Rate Variability. *Ergonomics,* Vol. 16, pp. 85–97.
[24] LUDWIG, C. (1847): Beiträge zur Kenntnis des Einflusses der Respirationsbewegungen auf den Blutumlauf im Aortensystem. *Arch. Anat. Physiol.,* pp. 242–302.
[25] MCCABE, P. M. – PORGES, S. W. – CARTER, C. S. (1981): Heart Period Variability During Estrogen Exposure and Withdrawal in Female Rats. *Physiology and Behavior,* Vol. 26, pp. 535–538.
[26] MULDER, G. – MULDER–HAJONIDES van der MEULEN, W. R. E. H. (1973): Mental Load and the Measurement of Heart Rate Variability. *Ergonomics,* Vol. 16, pp. 69–83.
[27] MULDER, G. (1980): The Heart of Mental Effort, Ph.D. Thesis, University of Groningen.
[28] MULDER, G. – MULDER, L. J. M. (1981): Information Processing and Cardiovascular Control. *Psychophysiology,* Vol. 14, p. 392.
[29] MULDER, L. J. M. (1988): Assessment of Cardiovascular Reactivity by Means of Spectral Analysis. Thesis. Groningen: University of Groningen.
[30] PAGANI, M. – LOMBARDI, F. – GUZZETTI, S. – RIMOLDI, O. – FURLAN, R. – PIZZINELLI, P. – SANDRONE, G. – MALFATTO, G. – DELL'ORTO, S. – PICCALUGA, E. – TURIEL, M. – BASELLI, G. – CERUTTI, S. – MALLIANI, A. (1986): Power Spectral Analysis of Heart Rate and Arterial Pressure Variabilities as a Marker of Sympathovagal Interaction in Man and Conscious Dog. *Circ. Res.,* Vol. 59, pp. 178–193.
[31] POMERANZ, B. – MACAULAY, R. J. B. – CAUDILL, M. A. – KUTZ, I. – ADAM, D. – GORDON, D. – KILBORN, K. M. – BARGER, A. C. – SHANNON, D. C. – COHEN, R. J. – BENSON, H. (1985): Assessment of Autonomic Functions in Humans by Heart Rate Spectral Analysis. *Am. J. Physiol.,* 248; (Heart and Circ. Physiol., 17), H151-H153.
[32] PORGES, S. W. – BOHRER, R. E. – KEREN, G.– CHEUNG, M. N. – FRANKS, G. J. – DRASGOW, F. (1981): The Influence of Methylphenidate on Spontaneous Autonomic Activity and Behavior in Children Diagnosed as Hyperactive. *Psychophysiol.,* Vol. 18, pp. 42–48.
[33] ROHMERT, W. – LAURIG, W. – PHILLIP, U. – LUCZAK, H. (1973): Heart Rate Variability and Work Load Measurement. *Ergonomics,* Vol. 16, pp. 33–44.
[34] SAUL, J. P. – REA, R. F. – ECKBERG, D. L. – BERGER, R. D. – COHEN, R.J. (1990): Heart Rate and Muscle Sympathetic Nerve Variability During Reflex Changes of Autonomic Activity. *Am. J. Physiol.,* Vol. 258, (Heart Circ. Physiol. 27), H713-H721.
[35] SAYERS, B. MCA., (1971): The Analysis of Cardiac Interbeat Interval Sequences and the Effect of Mental Work Load. *Proceedings of the Royal Society for Medicine,* Vol. 64, pp. 707–710.
[36] SAYERS, B. MCA. (1973): Analysis of Heart Rate Variability. *Ergonomics,* Vol. 16, pp. 17–32.
[37] SEKIGUCHI, C. – HANDA, J. – GOTOH, M. – KURIHARA, J. – NAGASAWA, J. – KURODA, I, (1979): Frequency Analysis of Heart Rate Variability Under Flight Conditions. *Aviation, Space and Environmental Medicine,* Vol. 50 (6) pp. 625–634.
[38] WEISE, F. – HEYDENREICH, F. – RUNGE, U. (1987): Contributions of Sympathetic and Vagal Mechanisms to the Genesis of Heart Rate Fluctuations During Orthostatic Load: a Spectral Analysis. *J. Auton. Nerv. Syst.,* Vol. 21, pp. 127–134.
[39] WESSELING, K. H. – SETTELS, J. J. (1985): Baromodulation Explains Short-term Blood Pressure Variability. In J. F. Orlebeke, G. Mulder and L .P. J van Dornen (Eds), The Psychophysiology of Cardiovascular Control (pp. 69–97). New York: Plenum Press.
[40] WIETHOFF, M. (1997): Task Analysis is Heart Work. PhD thesis. Delft University of Technology. Delft University Press.
[41] WIETHOFF, M. – IZSÓ, L. (1998) Mental Efficiency from the Point of View of Self-Regulation. *24th International Congress of Applied Psychology, Symposium: Sociopolitical Issues in Stress Research.* Program Book p. 60, August 9-14, 1998, San Francisco, U.S.A.
[42] WOMACK, B. F. (1971): The Analysis of Respiratory Sinus Arhythmia Using Spectral Analysis and Digital Filtering. *IEEE Trans. Biomed. Eng.,* Vol. 18, pp. 399–409.
[43] ZIJLSTRA, F. R. H. (1993): Efficiency in Work Behavior. PhD Thesis. Delft University of Technology. Delft University Press.