NIH Public Access

Author Manuscript

Clin Neurophysiol. Author manuscript; available in PMC 2012 October 1.

Published in final edited form as:

Clin Neurophysiol. 2011 October ; 122(10): 2059–2066. doi:10.1016/j.clinph.2011.02.032.

Sensitivity to Mental Effort and Test-Retest Reliability of Heart Rate Variability Measures in Healthy Seniors

Shalini Mukherjee[1], **Rajeev Yadav**[2], **Iris Yung**[3], **Daniel P. Zajdel**[1], and **Barry S. Oken**[1]

[1] Department of Neurology, Oregon Health & Science University, Portland, OR, USA

[2] Centre for Signal Processing and Communication (CENSIPCOM), Department of Electrical and Computer Engineering, Concordia University, Montreal, PQ, Canada

[3] School of Medicine, Oregon Health & Science University, Portland, OR, USA

Abstract

Objectives—To determine 1) whether heart rate variability (HRV) was a sensitive and reliable measure in mental effort tasks carried out by healthy seniors and 2) whether non-linear approaches to HRV analysis, in addition to traditional time and frequency domain approaches were useful to study such effects.

Methods—Forty healthy seniors performed two visual working memory tasks requiring different levels of mental effort, while ECG was recorded. They underwent the same tasks and recordings two weeks later. Traditional and 13 non-linear indices of HRV including Poincaré, entropy and detrended fluctuation analysis (DFA) were determined.

Results—Time domain (especially mean R-R interval/RRI), frequency domain and, among nonlinear parameters- Poincaré and DFA were the most reliable indices. Mean RRI, time domain and Poincaré were also the most sensitive to different mental effort task loads and had the largest effect size.

Conclusions—Overall, linear measures were the most sensitive and reliable indices to mental effort. In non-linear measures, Poincaré was the most reliable and sensitive, suggesting possible usefulness as an independent marker in cognitive function tasks in healthy seniors.

Significance—A large number of HRV parameters was both reliable as well as sensitive indices of mental effort, although the simple linear methods were the most sensitive.

Keywords

mental effort; working memory; heart rate variability; non-linear approaches; Poincaré; entropy; detrended fluctuation analysis

Corresponding Author: Barry Oken, Department of Neurology, Mail code CR120, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Portland, OR 97239, USA, Telephone number: 503-494-8873, Fax number: 503-494-9520, oken@ohsu.edu.

Introduction

Mental effort, a varying capacity for cognitive processing closely related to arousal and attention, has been studied using physiological measures (Kahneman, 1973; Kramer & Weber, 2000; Oken et al., 2006). One such measurement is heart rate variability (HRV), the analysis of beat-to-beat intervals. HRV is sensitive to several clinical conditions, especially those associated with cardiac disease or autonomic neuropathy (Task Force of the European Society of Cardiology, 1996; Hennessy, et al., 2001; Pope, et al., 2001; Sandercock, et al., 2005; Paul-Labrador, et al., 2006). Additionally, HRV has been used to assess interventions that might impact sympathetic-parasympathetic balance in conditions not associated with cardiac disease or autonomic neuropathy (Fu et al., 2006; Paul-Labrador, et al., 2006).

Previous research reported that HRV is sensitive to changes in mental effort (Mulder & Mulder, 1981; Aasman et al., 1987; Jorna, 1992; Veltman & Gaillard, 1993). With an increase in mental effort, there is a decrease in power around 0.10 Hz, in what they referred to as the mid-frequency band, 0.07–0.14 Hz (Jorna, 1992). This is now commonly known as the low frequency band, 0.04 – 0.15 Hz (Task Force of the European Society of Cardiology, 1996). The underlying mechanism may be due to increased sympathetic activation or the subjects' pattern of breathing (Althaus et al., 1998). In other studies, apart from a strong inverse relationship between mental effort and HRV power, mean RR interval appears to be the most sensitive measure (Capa et al., 2008; De Rivecourt et al., 2008; Henelius et al., 2009; Weippert et al., 2009).

While HRV has useful applications both clinically and as an autonomic measure in physiological research- the inter-subject variability of HRV measures is rather high (Gerritsen, et al., 2003; Pinna, et al., 2007). Thus, it is difficult to generalize the results from any one given study to others with subjects of different demographic characteristics. Additionally, the within-subject reliability of HRV measures remains unclear. The degree of reproducibility of HRV measurements from short-term recordings in healthy people is inconsistent and can at best be considered moderate (Pinna, et al., 2007; Sandercock, et al., 2005), except under highly controlled resting conditions (Melanson, 2000). Studies of reliability of HRV parameters during mental effort tasks is lacking in the literature.

Of late, non-linear methods of HRV analysis have gained prominence, with the promise of unraveling the inherent complexity in heart rate rhythms and an aim to model the complex interplay of influencing factors (Schumaker, 2004; González and Pereda, 2004). These new methods seem to contribute to the prognosis and treatment of different diseases (Huikuri et al., 2009). . While most studies using non-linear techniques have been carried out on long-term time series, like a 24-hour ambulatory Holter monitor (Bigger, et al., 1996; Huikuri, et al., 2000; Maestri, et al., 2007), it is of immense interest and benefit to see if these results are reproducible in short-term HRV, which is critical for studying cardiovascular autonomic changes. Shorter time series analysis techniques are desirable because many human psychophysiologic studies under controlled lab conditions are limited in how long they can be optimally maintained. While several methods suitable to assess short-term non-linear properties have been proposed (Guzzetti et al., 2005; Balocchi, et al., 2006), the reliability of these indices remains unclear. Studies carried out on patients with heart failure (Maestri, et al., 2006; Maestri, et al., 2007) and short term records in control subjects (Maestri, et al., 2007a) show a large variability in these indices Moreover, the reliability and usefulness of these indices as markers in the study of mental effort tasks has not been established. Hence, in the current study, we investigated the sensitivity and reliability of HRV to mental effort tasks in healthy seniors, using time domain, frequency domain and non-linear approaches to HRV analysis.

Methods

Participants

Forty participants from Portland, Oregon participated in the study. They were recruited for a clinical study on age-related cognitive function (Oken et al., 2008) and were compensated for their participation. The data for this analysis were obtained from testing sessions on two days separated by two weeks prior to the clinical study intervention. The Oregon Health & Science University Institutional Review Board approved the research, and each participant gave written informed consent.

Participants were screened and excluded if they had major medical conditions, or were taking medications that might interfere with the testing such as neuroleptics, beta-blockers, narcotic analgesics, and benzodiazepines. Other exclusions included scoring greater than 9 on the Center for Epidemiologic Studies Depression Scale (CESD) (Radloff, 1977), a history of insulin-dependent diabetes and significant liver, kidney or lung diseases. History of any cardiovascular condition such as congestive heart failure, myocardial infarction, arrhythmia, valvular disease, or uncontrolled hypertension (supine blood pressure greater than 160/95) was also among exclusion criteria. Participants with history of alcohol or recreational drug abuse, neurological diseases such as Parkinson's disease, stroke, dementia or a score greater than 6 on the Blessed Orientation-Memory-Concentration Test (Katzman et al., 1983) were excluded. We recorded their medications, dietary supplements, caffeine intake, alcohol intake and average sleep time to maximize consistency across visits. The test sessions were performed at the same time of day for all participants, at their preferred time either in the morning or afternoon. They were instructed to refrain from alcohol consumption for 24 hours prior to the testing but were allowed to ingest their usual dose of morning caffeine.

Materials

The working memory tasks occurred at least 60 minutes after the participants had been seated for the testing session. ECG was recorded while they performed two different versions of the working memory task: an 'easy' and a 'hard' trial. The task was adapted from the Sternberg memory task (Rypma & D'Esposito, 2000) that is sensitive to mental effort (Gevins et al, 1997) and has been used previously in the lab (Oken et al., 2008). Each block consisted of either all hard trials (six letters presented for a total of 3600 ms) or all easy trials (one letter presented for 1200 ms). Following a 3750 ms delay period after presentation of the letter strings, participants were shown a single letter probe stimulus and instructed to press one button if the probe was present in the memory sequence and another if it was not. The order of the task difficulty version was counter-balanced across visits. The easy and hard versions had different lengths so only the first 5 minutes of each were used for the HRV analysis. The median reaction times (RTs) to correctly answer the easy and hard versions trials were calculated.

Design

The two sessions were scheduled at the same time of day. Participants were seated in a dimly-lit, sound-attenuated room. As it was important for them to devote their entire attention to the task at hand during recordings, spontaneous breathing was used (i.e., there was no attempt to pace or otherwise control breathing). We present here the analysis of HRV data collected during the working memory task.

Recording Procedure

ECG from Lead I was digitized at 1000 Hz using BioSemi (BioSemi, Amsterdam, The Netherlands) amplifiers and a 150 Hz low-pass filter was applied. R-peaks were detected automatically and beat-to-beat intervals were extracted using Brain Vision Analyzer (Brain Products GmbH, Munich, Germany). One of the challenges encountered in HRV analysis is the handling of erroneous R-peaks in the ECG. Ectopy removal or replacement is shown to be essential regardless of the spectral estimation technique (Clifford & Tarassenko, 2005). Although the method used to handle erroneous RRIs have been shown to variously affect HRV estimates, a recent study, using pediatric and adult HR data showed that while the *identification* of outliers has an important impact, their *handling* (i.e., either dropping outliers or interpolation) had no impact on the estimation of HRV (Kemper et al., 2007). Time domain measures, like standard deviation of normal-to-normal intervals (SDNN) was least affected by editing methods (<5% error), although frequency domain measures were affected (Salo et al., 2001)

In our study, the data was first visually inspected and then processed using the signal quality indices and Kalman filter proposed by Li et al., in 2007. This method provides an accurate HR estimate even in the presence of high levels of persistent noise and artifact, and during episodes of extreme bradycardia and tachycardia. However, in our data set, less than 5% of the total number of RRIs were either erroneous detections/artifacts or aberrant beats (e.g., premature contractions) and were eliminated from the analysis. The true time stamp of annotated R-R files The RR intervals with their associated time stamps, as derived from the previous processing step, were then processed using MATLAB version 7.9 (MathWorks Inc., MA, USA) using and programs developed in-house.

HRV Analysis

Time domain—The common measures of HRV were computed: mean NN or RR intervals (RRI), SDNN (standard deviation of NN intervals), RMSSD (square root of the mean squared difference of the consecutive NN intervals), NN50 (the number of pairs of successive NNs that differ by more than 50 ms.) and pNN50 (the proportion of NN50 divided by total number of NNs).

Spectral analysis—This was carried out using three different methods: a Fast Fourier transform (FFT), Autoregressive (AR) and modified Lomb periodogram (MLP) analysis. For both AR and FFT, the annotated RR time series was cubic spline interpolated at 4 Hz. The FFT power spectral density (PSD) was calculated using the Welch method. The MLP analysis was carried out according to Thong et al.,2004 on the non-interpolated time series. This method allows spectral analysis directly on unevenly sampled data and therefore, no interpolation is required. PSD was estimated in two frequency bands of low frequency (LF) of 0.04 –0.15 Hz and high frequency (HF) of 0.15–0.4 Hz (Pomeranz, et al., 1985; Pumprla et al., 2002)

Non-linear analysis—The R-R interval (RRI) time series exhibit both irregularity and self-similarity. The irregularities in the RRI were quantified by entropy-based methods. As there are no set guidelines for non-linear analysis of HRV (Task Force of the European Society of Cardiology, 1996), investigators have used different methods and their exact role in autonomic modulation is still a matter of debate. Some studies have shown that a decrease in entropy points to a perturbation of the complex physiological mechanism and is a prognostic factor of diseases affecting the cardiovascular autonomic nervous system (Beckers et al., 2001; Huikuri, et al., 2009). However, we found no studies to indicate whether these indices are useful and reliable markers for mental effort. Therefore, we studied a variety of entropy-based methods to quantify mental effort. The non-linear measures considered in this study are as mentioned below:

- The *Poincaré* plot is the most commonly used non-linear estimator of HRV function. Each successive RRI is plotted against its previous RRI. After fitting an ellipse and plotting two axes (perpendicular to each other) to the points, one can calculate the standard deviation of the distance of the points from each axis. The SD1 (minor axis) value reflects short-term variability while SD2 (major axis) reflects long-term variability. Poincaré plots can be valid markers for mental stress (Cervantes et al., 2009).

- *Fractal dimension* (FD) is a geometric approach to measure self-similarity in a time series and is suited to measure complex fluctuations with statistical properties observed in RRI. FD was significantly reduced in schizophrenia patients compared to controls (Bar, et al., 2007) and was relatively independent of heart rate. In patients with stroke involving multiple cerebral vascular territories, FD was a useful measure to distinguish between different lesion severity and from controls (D'Addio, et al., 2009). It has also been seen to be a marker of sidedness of stroke (He et al., 2010).
- *Detrended fluctuation analysis (DFA) is* another measure to quantify the fractal scaling properties of the short interval RRI time series. The root-mean-square fluctuation of an integrated and detrended time-series is measured at different scales. The fluctuations are characterized by a scaling exponent which is the slope of the line relating to log of the fluctuation on a log scale. DFA is significantly diminished in participants with major depressive disorder when compared to controls (Schulz et al., 2010).
- *Lyapunov exponent* is a simple non-linear measure of how fast two initially nearby points on a trajectory will diverge from each other as the system evolves, and a positive Lyapunov exponent is a strong indicator of chaos. Even though an *m* dimensional system has *m* Lyapunov exponents, the *largest Lyapunov exponent* (LLE) is sufficient to represent the whole system. Recent studies have shown LLE is good indicator of anxiety and mental stress and is sensitive to the effect of cardio-inhibitory tricyclic antidepressants (Yeragani et al., 2004; Yeragani & Rao, 2003).
- *Entropy* measures express the degree of randomness in the cardiovascular system. In healthy people, the tachogram would show high entropy. In the case of low entropy systems, the tachogram would be ordered and repetitive, showing low variability. *Spectral entropy* is a measure of uncertainty about the event at a given frequency and a direct measurement of system complexity. *Sample entropy* is a regularity statistic that quantifies the unpredictability of fluctuations in the RRI time series. *Multiscale entropy* aims to separate short-range and long-range correlations in the signal. Often entropy increases with the degree of irregularity and is greatest for completely random systems and may not always be associated with an increase in the complexity. This is possible because the entropy measure is a single scale analysis, and therefore, cannot account for temporal fluctuations of complex healthy systems. Entropy decreases in various medical conditions like schizophrenia (Bar, et al., 2007), major depressive disorder (Schulz, et al., 2010), cardiovascular diseases, and autonomic neuropathy (Khandoker et al., 2010). *Lempel-Ziv complexity* (LZ) is another entropy metric used to measure the inherent complexity of discrete time physiologic signals. A more dynamic and robust system has a higher complexity measure.
- The Spectral *skewness* and spectral *kurtosis* provides important information regarding the distribution and shape of the power spectral density.

Statistical analysis

The LF power was normalized by the formula: LF/(LF + HF) * 100. The HF power was normalized in a similar manner and they are denoted as normalized percentages (np) in Tables 1 and 2. Distribution of the data-set was estimated by the Kolmogorov-Smirnov tests for normality. Since most of the variables were not normally distributed, non-parametric analyses were performed to enable meaningful comparisons. Test-retest reliability of HRV parameters was estimated using the *Kendall's tau* (τ) statistic. This has been found to be a more suitable test to measure non-parametric correlation than the Spearman's rho and provides more reliable confidence intervals (Kendall, 1990; Newson, 2002).

Visual assessment of reliability was carried out using Bland-Altman plots (Bland & Altman, 1986). A plot of the average of two measurements by the difference among two measurements provides a qualitative estimate of absolute reliability. Absolute reliability is a measure of the degree of variability on repeated measurements of a variable, in an individual. Mean differences should be closer to zero to demonstrate agreement. The measured values should lie within the limits of agreement, i.e. mean difference ±1.96*SD of the mean difference. In the case of a reliable measure, 95% of the value points should lie between these limits. The coefficient of repeatability was also calculated (Atkinson & Nevill, 1998; Bland & Altman, 1999). This is a precision measure which represents the value below which the absolute difference between two repeated test results occurs with a probability of 95%. Hence, smaller values denote a higher concordance in intra-individual scores on repeated measurements. After plotting the graphs, they were visually inspected for heteroscedasticity (i.e., if there is more of a variability or scatter at the higher end of the values). In the cases where visual inspection revealed a "funnel-shaped" association between the mean and the difference of a measure, it was log-transformed and re-plotted. Values were back transformed before reporting the mean difference and confidence intervals (Bland & Altman, 1999).

In order to find the effect of gender and age on the sensitivity of HRV measures, we carried out a general linear model (GLM) multivariate analysis, with gender as a fixed effect and age as a covariate, in the model. The age of the participants was categorized into two groups; those below or above the median age of 73 years. Non-normal variables were log transformed for the analysis and rechecked for normality, using Kolmogorov-Smirnov tests. In order to estimate the sensitivity of HRV variables to the difficulty of the mental effort tasks, a non-parametric Wilcoxon's signed rank test was carried out. The effect size of the HRV measures was estimated using the Goodman and Kruskal's gamma statistic. This association ranges from −1 to +1. Gamma has been shown to be the nonparametric analog to Cohen's d (Hogarty & Kromrey, 2000).

We used SPSS V.18 (SPSS Inc., IL, USA) for statistical analysis and MedCalc V.11.1.1 (MedCalc Software, Belgium) for calculating the Bland-Altman plots and obtain confidence intervals for Kendall's τ measures. A two-sided *p*-value, $p < 0.05$, was considered significant for all statistical tests. For the sensitivity analysis, given the presence of 21 HRV measures, a simple Bonferroni correction was employed. A p-value <0.002 would be definitely significant and p-values between 0.002 and 0.05 would be of uncertain significance.

Results

Data were obtained from 40 participants. One participant was excluded from Visit 1 and 6 from Visit 2, due to arrhythmia or signal artifacts.

Demographic characteristics

Mean age of the participants was 73.05 ± 4.92 years, with the maximum and minimum ages being 83 and 65 respectively. There were 26 females and 14 males. Mean age in females was 72.81 ± 4.65 years while that in males was 73.5 ± 5.54 years.

Analysis of memory task scores

The reaction times (RTs) for the modified Sternberg tasks were: Visit 1– 846.14 ± 204.69ms (easy trial) and 1135.74 ± 171.19ms (hard trial); Visit 2– 808.41 ± 175.19ms (easy trial) and 1064.25 ± 182.83ms (hard trial) respectively. The RTs for the easy and hard tasks were different ($p<0.0001$) for both visits, demonstrating the significant effect of task difficulty. The RT differences between the easy and hard version of the Sternberg working memory task, a measure of task difficulty, during Visit 1, was 291.61 ± 123.79ms and during Visit 2 was 247.2 ± 121.95ms. They were highly correlated (Kendall's τ of 0.528, $p <0.0001$). Of some note, subjects had significantly shorter reaction times on the easy and hard versions as well as the differences on the second visit (p 0.004).

Reliability of HRV measures

The tests for reliability of HRV measures are tabulated in Table 1. Spectral analysis of HRV was carried out using three methods: Autoregressive, FFT and a Lomb periodogram. Of these, the Lomb method emerged as the most reliable and sensitive method (data not shown) and hence was finally considered for presentation and analysis. Analyzing the Kendall's τ values for test-retest reliability of the HRV parameters, we found that 17/21 parameters (81%) were significant for the easy and hard trials for both visits (Table 1). Time domain measures were the most reliable in both trial conditions; Kendall's τ for time domain measures ranged from 0.74 to 0.26 (*p*-values from <0.001 to 0.03) for the easy trial and 0.74 to 0.29 (*p*-value <0.0001 to 0.02) for the hard trial. This was followed by frequency estimates by Lomb periodogram (0.36 to 0.31, *p*-values 0.003 to 0.01) for the easy and 0.38 to 0.36 (*p*-values 0.001 to 0.003) for the hard trial. For non-linear variables, the number which were significantly reliable (i.e. 8/13 variables) by Kendall's τ estimates, similar for the easy and hard trials (8/13 scores were significant). Among them, Poincaré (easy trial: from 0.39 to 0.25, *p*-value 0.001 to 0.04; hard trial: from 0.35 to 0.29, *p*-value 0.005 to 0.02) were the most reliable. The only other non-linear measures that were consistently reliable in both the easy and hard trials were DFA-α1 (τ = 0.36 and 0.35, *p*-values 0.006 and 0.005), multiscale entropy (τ = 0.26 and 0.32, *p*-values 0.04 and 0.01) and spectral skewness (τ = 0.28 and *p*-value 0.02 for both) in the easy and hard trials respectively.

For absolute reliability, the difference of means and coefficient of repeatability (CR) were smallest for the time domain measures (e.g. CR-RMSSD was −0.002 and −0.001 for easy and hard trials respectively). NN50 and pNN50 had a higher difference of means and CR, demonstrating a lower absolute reliability, although they had higher Kendall τ values (Table 1). Using an arbitrary cut-off value of 1 for CR, we saw that LF, and especially HF were reliable indicators (CR for log transformed HF were 0.61 and 0.5 for easy and hard trials) for mental effort task. These also showed a higher concordance with obtained mean difference values. Absolute reliability estimates (i.e. low mean differences and low CR scores) were high for many of the non-linear parameters. DFA was especially reliable for both easy and hard trials. Eleven out of the 13 parameters studied had CR <1 in the easy and hard trials and lower corresponding mean differences, which meant they were reliable measures.

Sensitivity of HRV measures

A multivariate analysis showed that neither gender nor age had a significant impact on group differences (i.e. did not meet the Bonferroni criteria of $p<0.0002$); hence they were not considered in the analysis. HRV measures were compared among the easy and hard trials to estimate whether they were sensitive to the mental task difficulty. They emerged as being comparably sensitive on both visits (10/21 variables on Visit 1 and 6/21 variables on Visit 2). The time domain measures were the most sensitive, with mean RR, SDNN, RMSSD achieving significance ($p< 0.0001$). None of the frequency domain measures achieved significance with a Bonferroni adjustment. Among the non-linear methods, only Poincaré was sensitive enough to detect significant differences after Bonferroni adjustment ($p<0.0001$ on both visits, Table 2).

In order to better understand the usefulness of HRV measures in mental effort tasks, we also further considered the effect size of the observed relationship. Hence, a Goodman and Kruskall's gamma statistic was computed for all the variables. This in turn, would be very helpful in selecting HRV parameters, which could be used in future experiments. Most of the parameters had significant effect sizes (20/21 parameters, $p<0.005$) for Visits 1 and 2, although the absolute values of the gamma statistic were not high. The highest gamma values were for mean RRI (0.91 for Visit 1 and 0.89 for Visit 2). Hence, considering an arbitrary cutoff of Gamma =0.5, we observed that most time domain, Lomb frequency domain and Poincaré measures had large effect sizes on both Visits 1 and 2. Among non-linear indices, DFA α, α1, sample entropy, kurtosis and skewness had large effect sizes during Visit 2.

Discussion

Although HRV is widely used as a physiological measure in cognitive function studies, the reliability of these measures during cognitive tasks is not known. In our study, most of the HRV parameters were reliable. Test-retest reliability (Kendall's τ) closely correlated to the absolute reliability measures using mean difference scores and coefficient of repeatability, over both easy and hard trials. Among the most reliable- time domain > frequency domain >Poincaré estimates. Among other non-linear measures, DFA-α1 and multiscale entropy were reliable over both task conditions. Since DFA-α1 is said to closely correlate to: LF/(LF + HF) (Francis et al. 2002), it could be a reflection of the highly reliable frequency domain measures.

We found the Lomb method to be the most reliable spectral estimate, when compared to FFT and AR. Previous researchers have proposed PSD estimates using the Lomb-Scargle periodogram to be a more appropriate method for an unevenly sampled signal, such as RR-interval data (Laguna, Moody, & Mark, 1998; Lomb, 1976; Scargle, 1982). It outperforms both AR and FFT methods in terms of accuracy in irregularly-placed data like in inter-beat intervals of heart rate (Chang, et al., 2001; Moody, 1993). The re-sampling process in FFT tends to overestimate the LF/HF ratio in HRV, which adds to the LF component and reduces the HF content. The Lomb periodogram on the other hand generates a more accurate and less noisy estimation of LF/HF ratios (Clifford & Tarassenko, 2005).

As mentioned above, we observed RRI and the time domain measures to be highly reliable outcome measures over the task conditions. This suggests that heart rate (HR) itself can be considered a strong and reliable outcome measure in mental effort tasks. Mean HR, which can be easily recorded in a study (i.e., HR= 60/RRI), has previously been demonstrated to be more reproducible than spectral power estimates during incremental head-up tilt (Bootsma et al., 1996). Time domain metrics of RRI were also found to be the most sensitive measure of HRV in cognitive function tasks (Henelius, et al., 2009). Hence, cognitive functional changes associated with mental effort task likely represents a different autonomic profile than cardiovascular disorders, in which PSD and non-linear analysis can provide much more information than only HR (Akselrod et al., 1981; Stein et al., 2005). However, they do provide more information that HR alone; we can identify the sympathetic and parasympathetic cardiac drives and the inherent complexity of the HR signal using these measures.

The degree of task difficulty was evident from the difference in working memory RTs in the easy and hard trials, which was a reliable indicator over the two visits. Again, time domain measures were the most sensitive to differentiate between the two trials over both visits. The LF/HF ratio, generally regarded as a measure of sympatho-vagal balance, was significantly higher in hard trials during the first visit, but failed to show any difference during Visit 2. HF fluctuations have been attributed to vagal modulation and LF fluctuations appear to be jointly mediated by sympathetic and vagal influences, probably through the baroreflex mechanism (Aubert & Ramaekers, 1999). Among non-linear indices, Poincaré and DFA-$\alpha 2$ were the most sensitive and had larger effect sizes; although in the case of DFA, it was evident only in Visit 1. Spectral entropy was greater for the easy trial during visit 2, probably reflecting the influence of the task difficulty on that measure. A direct correlation was observed between the sensitivity and effect size of the time domain, LF/HF ratio and Poincaré measures. Most of the other measures did not distinguish between the trials, regardless of their significant effect sizes. The difference in sensitivity was likely due to the inherent variability in HRV measures. The exact implications of many of these non-linear measures with respect to mental effort in unknown; they have not been explored previously.

There are many factors which might alter the reproducibility and sensitivity of HRV. A small sample size, gender, circadian rhythm, age of participants (which is associated with decrease in total power) and breathing condition (paced versus spontaneous) are some of them. Other factors include body position, participants activity level prior to the recording (i.e., how long the participants have been inactive) or the lack of some normalizing transform (either relative power or log power), which decreases measures of variability compared to absolute power. In our study, the LF and HF powers were normalized as percentage of LF + HF. We also took special precautions to maintain similar time of day and activity level between visits in all participants. In our study, participants had been seated for at least 60 minutes prior to ECG recording in order to remove the confounding influence of previous activity.

The current study faced certain limitations. The already lengthy cognitive testing session (approximately 90-min duration), precluded allocating an additional 5-minute block for a baseline recording while not performing any task. Controlled breathing was not used because asking participants to concentrate on their breathing while also being asked to simultaneously attend to a cognitive task would be difficult. This would have also changed the task by introducing dual-task demands. The effect of breathing pattern on HRV is debatable. Some studies demonstrate that different breathing conditions may have an impact on the reproducibility of HRV (Pitzalis, et al., 1996; Pinna et al., 2007). In contrast, other studies found that such factors did not have a significant impact on HRV reliability (Sinnreich et al., 1998; Maestri et al., 2010). The findings of the latter seem to suggest that HRV is reliable and consistent over time, whether or not respiration is controlled.

The age group in our study was selective, with all participants in the study between the ages of 65 to 83. This was in part to enable the examination of cognitive function in this older age group and its relationship with HRV but also due to the fact that HRV tends to vary significantly across age. Therefore studying a homogenous group in terms of age range allows for more consistency and reliability. In our study, we found that neither age nor gender significantly influenced the HRV measures (after Bonferroni correction). This is likely due to the fact that both these factors play a more important role in a younger population and these differences diminish with age. Studies have found a stabilization in age related decline in both linear (Antelmi et al., 2004) and non-linear HRV (Beckers, Verhayden & Aubert, 2006) parameters after the fourth decade. The latter study on healthy seniors also couldn't establish any clear differences between sexes. However, because of the use of senior subjects and the prevalence of age-associated medical conditions impacting HRV, it may be difficult to generalize these results to a broader population.

The current study finds heart rate and certain HRV measures to be a reliable and sensitive tool in the evaluation of autonomic changes accompanying mental effort. Mean RRI, time domain, frequency domain (especially the Lomb periodogram) and Poincaré were the most reliable and sensitive HRV parameters although there is a suggestion that the simpler linear methods may have been most useful for assessing mental effort over short time periods. Many of the non-linear analysis measures were reliable, but not sensitive enough or have an adequate effect size in order to qualify as useful indices of mental effort in our study. These include DFA-α, fractal dimension, LLE, sample entropy and kurtosis. DFA-α2 and spectral entropy were sensitive to the task difficulty during Visits 1 and 2 respectively. Their effect sizes, though statistically significant, were not large. These results may help guide the choice of HRV parameters to assess change associated with cognitive performance in senior subjects in future studies.

Acknowledgments

This study was supported by NIH U19 AT002656 and K24 AT005121. Roger Ellingson, M.S. provided software support for the heart rate variability data analysis and Ms. Alexandra Amen helped with arranging the references.

References

Aasman J, Mulder G, Mulder LJ. Operator effort and the measurement of heart-rate variability. Hum Factors. 1987; 29:161–170. [PubMed: 3610181]

Akselrod S, Gordon D, Ubel FA, Shannon DC, Berger AC, Cohen RJ. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. Science. 1981; 213:220–222. [PubMed: 6166045]

Althaus M, Mulder LJ, Mulder G, Van Roon AM, Minderaa RB. Influence of respiratory activity on the cardiac response pattern to mental effort. Psychophysiology. 1998; 35:420–430. [PubMed: 9643056]

Antelmi I, de Paula RS, Shinzato AR, Peres CA, Mansur AJ, Grupi CJ. Influence of age, gender, body mass index, and functional capacity on heart rate variability in a cohort of subjects without heart disease. Am J Cardiol. 2005; 93:381–385. [PubMed: 14759400]

Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. Sports Med. 1998; 26:217–238. [PubMed: 9820922]

Aubert AE, Ramaekers D. Neurocardiology: the benefits of irregularity. The basics of methodology, physiology and current clinical applications. Acta Cardiol. 1999; 54:107–120. [PubMed: 10478266]

Balocchi R, Cantini F, Varanini M, Raimondi G, Legramante JM, Macerata A. Revisiting the potential of time-domain indexes in short-term HRV analysis. Biomed Tech (Berl). 2006; 51:190–193. [PubMed: 17061936]

Bar KJ, Boettger MK, Koschke M, Schulz S, Chokka P, Yeragani VK, et al. Non-linear complexity measures of heart rate variability in acute schizophrenia. Clin Neurophysiol. 2007; 118:2009–2015. [PubMed: 17646130]

Beckers F, Ramaekers D, Aubert AE. Approximate Entropy of Heart Rate Variability: Validation of Methods and Application in Heart Failure. Cardiovasc Eng. 2001; 1:177–182.

Beckers F, Varhayden B, Aubert AE. Aging and non linear heart rate control in healthy population. Am J Physiol Heart Circ Physiol. 2006; 290:H2560–H2570. [PubMed: 16373585]

Bigger JT Jr, Steinman RC, Rolnitzky LM, Fleiss JL, Albrecht P, Cohen RJ. Power law behavior of RR-interval variability in healthy middle-aged persons, patients with recent acute myocardial infarction, and patients with heart transplants. Circulation. 1996; 93:2142–2151. [PubMed: 8925583]

Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res. 1999; 8:135–160. [PubMed: 10501650]

Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet. 1986; 1:307–310. [PubMed: 2868172]

Bootsma M, Swenne CA, Bruschke AV. Heart rate variability during repeated incremental head-up tilt discloses time dependence of individual autonomic dynamics. Clin Cardiol. 1996; 19:62–68. [PubMed: 8903540]

Capa RL, Audiffren M, Ragot S. The interactive effect of achievement motivation and task difficulty on mental effort. Int J Psychophysiol. 2008; 70:144–150. [PubMed: 18627781]

Cervantes Blasquez JC, Rodas Font G, Capdevila Ortis L. Heart-rate variability and precompetitive anxiety in swimmers. Psicothema. 2009; 21:531–536. [PubMed: 19861094]

Chang KL, Monahan KJ, Griffin MP, Lake D, Moorman JR. Comparison and clinical application of frequency domain methods in analysis of neonatal heart rate time series. Ann Biomed Eng. 2001; 29:764–774. [PubMed: 11599584]

Clifford GD, Tarassenko L. Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. IEEE Trans Biomed Eng. 2005; 52:630–638. [PubMed: 15825865]

D'Addio G, Corbi G, Accardo A, Russo G, Ferrara N, Mazzoleni MC, et al. Fractal behaviour of heart rate variability reflects severity in stroke patients. Stud Health Technol Inform. 2009; 150:794–798. [PubMed: 19745422]

De Rivecourt M, Kuperus MN, Post WJ, Mulder LJ. Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight. Ergonomics. 2008; 51:1295–1319. [PubMed: 18802817]

Francis DP, Willson K, Georgiadou P, Wensel R, Davies LC, Coats A, et al. Physiological basis of fractal complexity properties of heart rate variability in man. J Physiol. 2002; 542(Pt 2):619–629. [PubMed: 12122157]

Fu CH, Yang CC, Lin CL, Kuo TB. Effects of long-term vegetarian diets on cardiovascular autonomic functions in healthy postmenopausal women. Am J Cardiol. 2006; 97:380–383. [PubMed: 16442400]

Gerritsen J, Tenvoorde B, Dekker J, Kingma R, Kostense P, Bouter L, et al. Measures of cardiovascular autonomic nervous function: agreement, reproducibility, and reference values in middle age and elderly subjects. Diabetologia. 2003; 46:330–338. [PubMed: 12687330]

Gevins A, Smith ME, McEvoy L, Yu D. High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cereb Cortex. 1997; 7:374–385. [PubMed: 9177767]

Gonzalez JJ, Pereda E. Applications of fractal and non-linear time series analysis to the study of short-term cardiovascular control. Curr Vasc Pharmacol. 2004; 2:149–162. [PubMed: 15320516]

Guzzetti S, Borroni E, Garbelli PE, Ceriani E, Della Bella P, Montano N, et al. Symbolic dynamics of heart rate variability: a probe to investigate cardiac autonomic modulation. Circulation. 2005; 112:465–470. [PubMed: 16027252]

He L, Li C, Luo Y, Dong W, Yang H. Clinical prognostic significance of heart abnormality and heart rate variability in patients with stroke. Neurol Res. 2010; 32:530–534. [PubMed: 19473556]

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Circulation. 1996; 93:1043–1065. [PubMed: 8598068]

Henelius A, Hirvonen K, Holm A, Korpela J, Muller K. Mental workload classification using heart rate metrics. Conf Proc IEEE Eng Med Biol Soc. 2009; 2009:1836–1839. [PubMed: 19963519]

Hennessy MJ, Tighe MG, Binnie CD, Nashef L. Sudden withdrawal of carbamazepine increases cardiac sympathetic activity in sleep. Neurology. 2001; 57:1650–1654. [PubMed: 11706106]

Hogarty, KY.; Kromrey, JD. Robust Effect Size Estimates and Meta-Analysis Tests of Homogeneity. Proceedings of SAS Users Group International. 2000(Proceedings of SAS Users Group International); p. 1129-1144.

Huikuri HV, Makikallio TH, Peng CK, Goldberger AL, Hintze U, Moller M. Fractal correlation properties of R-R interval dynamics and mortality in patients with depressed left ventricular function after an acute myocardial infarction. Circulation. 2000; 101:47–53. [PubMed: 10618303]

Huikuri HV, Perkiomaki JS, Maestri R, Pinna GD. Clinical impact of evaluation of cardiovascular control by novel methods of heart rate dynamics. Philos Transact A Math Phys Eng Sci. 2009; 367:1223–1238. [PubMed: 19324705]

Jorna PG. Spectral analysis of heart rate and psychological state: a review of its validity as a workload index. Biol Psychol. 1992; 34:237–257. [PubMed: 1467395]

Kahneman, D., editor. Attention and Effort. Englewood Cliffs, New Jersey: Prentice-Hall, Inc; 1973.

Katzman R, Brown T, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. Am J Psychiatry. 1983; 140:734–739. [PubMed: 6846631]

Kemper KJ, Hamilton C, Atkinson M. Heart rate variability: impact of differences in outlier identification and management strategies on common measures in three clinical populations. Pediatr Res. 2007; 62:337–42. [PubMed: 17597640]

Kendall, MGGJ., editor. Rank Correlation Methods. 5. London: Arnold; 1990.

Khandoker AH, Jelinek HF, Moritani T, Palaniswami M. Association of cardiac autonomic neuropathy with alteration of sympatho-vagal balance through heart rate variability analysis. Med Eng Phys. 2010; 32:161–167. [PubMed: 20004128]

Kramer, A.; Weber, T., editors. Applications of psychophysiology to human factors. New York: Cambridge University Press; 2000.

Laguna P, Moody GB, Mark RG. Power spectral density of unevenly sampled data by least-square analysis: performance and application to heart rate signals. IEEE Trans Biomed Eng. 1998; 45:698–715. [PubMed: 9609935]

Li Q, Mark RG, Clifford GD. Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter. Physiol Meas. 2008; 29:15–32. [PubMed: 18175857]

Lomb NR. Least-squares frequency analysis of unequally spaced data. Astrophysics and Space Science. 1976; 39:447–462.

Maestri R, Pinna GD, Balocchi R, D'Addio G, Ferrario M, Porta A, et al. Clinical correlates of non-linear indices of heart rate variability in chronic heart failure patients. Biomed Tech (Berl). 2006; 51:220–223. [PubMed: 17061943]

Maestri R, Pinna GD, Accardo A, Allegrini P, Balocchi R, D'Addio G, et al. Nonlinear indices of heart rate variability in chronic heart failure patients: redundancy and comparative clinical value. J Cardiovasc Electrophysiol. 2007; 18:425–433. [PubMed: 17284264]

Maestri R, Pinna GD, Porta A, Balocchi R, Sassi R, Signorini MG, et al. Assessing nonlinear properties of heart rate variability from short-term recordings: are these measurements reliable? Physiol Meas. 2007a; 28:1067–1077. [PubMed: 17827654]

Maestri R, Raczak G, Danilowicz-Symanowicz L, Torunski A, Sukiennik A, Kubica J, et al. Reliability of heart rate variability measurements in patients with a history of myocardial infarction. Clin Sci (Lond). 2010; 118:195–201. [PubMed: 20310082]

Melanson EL. Resting heart rate variability in men varying in habitual physical activity. Med Sci Sports Exerc. 2000; 32:1894–1901. [PubMed: 11079519]

Moody GB. Spectral analysis of heart rate without resampling. Comput Cardiol. 1993; 20:715–718.

Mulder G, Mulder LJ. Information processing and cardiovascular control. Psychophysiology. 1981; 18:392–402. [PubMed: 7267921]

Newson R. Parameters behind "nonparametric" statistics: Kendall's tau, Somers' D and median differences. The Stata Journal. 2002; 2:45–64.

Oken BS, Flegal K, Zajdel D, Kishiyama S, Haas M, Peters D. Expectancy effect: impact of pill administration on cognitive performance in healthy seniors. J Clin Exp Neuropsychol. 2008; 30:7–17. [PubMed: 18165936]

Oken BS, Salinsky MC, Elsas SM. Vigilance, alertness, or sustained attention: physiological basis and measurement. Clin Neurophysiol. 2006; 117(9):1885–1901. [PubMed: 16581292]

Paul-Labrador M, Polk D, Dwyer JH, Velasquez I, Nidich S, Rainforth M, et al. Effects of a randomized controlled trial of transcendental meditation on components of the metabolic syndrome in subjects with coronary heart disease. Arch Intern Med. 2006; 166:1218–1224. [PubMed: 16772250]

Pinna GD, Maestri R, Torunski A, Danilowicz-Szymanowicz L, Szwoch M, La Rovere MT, et al. Heart rate variability measures: a fresh look at reliability. Clin Sci (Lond). 2007; 113:131–140. [PubMed: 17381425]

Pitzalis MV, Mastropasqua F, Massari F, Forleo C, Di Maggio M, Passantino A, et al. Short- and long-term reproducibility of time and frequency domain heart rate variability measurements in normal subjects. Cardiovasc Res. 1996; 32:226–233. [PubMed: 8796108]

Pomeranz B, Macaulay RJ, Caudill MA, Kutz I, Adam D, Gordon D, et al. Assessment of autonomic function in humans by heart rate spectral analysis. Am J Physiol. 1985; 248(1 Pt 2):H151–153. [PubMed: 3970172]

Pope CA 3rd, Eatough DJ, Gold DR, Pang Y, Nielsen KR, Nath P, et al. Acute exposure to environmental tobacco smoke and heart rate variability. Environ Health Perspect. 2001; 109:711–716. [PubMed: 11485870]

Pumprla J, Howorka K, Groves D, Chester M, Nolan J. Functional assessment of heart rate variability: physiological basis and practical applications. Int J Cardiol. 2002; 84:1–14. [PubMed: 12104056]

Radloff L. The CES-D scale: a self-report depression scale for research in the general population. Applied Psychological Measurement. 1977; 1:385–401.

Rypma B, D'Esposito M. Isolating the neural mechanisms of age-related changes in human working memory. Nat Neurosci. 2000; 3:509–515. [PubMed: 10769393]

Salo MA, Huikuri HV, Seppänen T. Ectopic beats in heart rate variability analysis: effects of editing on time and frequency domain measures. Ann Noninv Electrocardiol. 2001; 6:5–17.

Sandercock GR, Bromley PD, Brodie DA. The reliability of short-term measurements of heart rate variability. Int J Cardiol. 2005; 103:238–247. [PubMed: 16098384]

Scargle JD. Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. Astrophysical Journal. 1982; 263:835–853.

Schulz S, Koschke M, Bar KJ, Voss A. The altered complexity of cardiovascular regulation in depressed patients. Physiol Meas. 2010; 31:303–321. [PubMed: 20086275]

Schumacher A. Linear and nonlinear approaches to the analysis of R-R interval variability. Biol Res Nurs. 2004; 5:211–221. [PubMed: 14737922]

Sinnreich R, Kark JD, Friedlander Y, Sapoznikov D, Luria MH. Five minute recordings of heart rate variability for population studies: repeatability and age-sex characteristics. Heart. 1998; 80:156–162. [PubMed: 9813562]

Stein PK, Domitrovich PP, Huikuri HV, Kleiger RE. Traditional and nonlinear heart rate variability are each independently associated with mortality after myocardial infarction. J Cardiovasc Electrophysiol. 2005; 16:13–20. [PubMed: 15673380]

Thong, T.; Yung, I.; Zajdel, D.; Ellingson, R.; McNames, J.; Aboy, M., et al., editors. Heart rate variability analysis of effect of nicotine using periodograms. Conference Proceedings, 26th Annual international Conference of the IEEE Engineering in Medicine and Biology Society; 2004. p. 294-297.

Toyry J, Mantysaari M, Hartikainen J, Lansimies E. Day-to-day variability of cardiac autonomic regulation parameters in normal subjects. Clin Physiol. 1995; 15:39–46. [PubMed: 7712691]

Veltman JA, Gaillard AW. Indices of mental workload in a complex task environment. Neuropsychobiology. 1993; 28:72–75. [PubMed: 8255413]

Weippert M, Thielmann B, Stoll R, Pfister EA, Bockelmann I. [Sympatho-vagal balance and cardiac response to mental challenge]. Biomed Tech (Berl). 2009; 54:197–203. [PubMed: 19614267]

Yeragani VK, Mallavarapu M, Radhakrishna RK, Tancer M, Uhde T. Linear and nonlinear measures of blood pressure variability: increased chaos of blood pressure time series in patients with panic disorder. Depress Anxiety. 2004; 19:85–95. [PubMed: 15022143]

Yeragani VK, Rao R. Effect of nortriptyline and paroxetine on measures of chaos of heart rate time series in patients with panic disorder. J Psychosom Res. 2003; 55:507–5. [PubMed: 14642980]

Table 1

Comparison of test-retest reliability of HRV for easy and hard trial sets over visits 1 and 2

	Reliability between Easy trial Visit 1 and Visit 2				Reliability between Hard trial, Visit 1 and 2			
	Difference of Means (95% CI)	**CR**	**Kendal's tau (95% CI)**	**p-value**	**Difference of Means (95% CI)**	**CR**	**Kendal's tau (95% CI)**	**p-value**
Mean RR	0.03 (0.0002, 0.05)	0.03	0.74 (0.6, 0.83)	<0.0001	0.02 (−0.003, 0.05)	0.02	0.74 (0.63, 0.84)	<0.0001
SDNN	0.003 (−0.006, 0.02)	0.01	0.26 (−0.02, 0.51)	0.03	0.002 (−0.002, 0.007)	0.008	0.29 (0.06, 0.51)	0.02
RMSSD	−0.0002 (−0.01, 0.01)	0.01	0.39 (0.13, 0.59)	0.001	−0.001 (−0.007, 0.005)	0.01	0.35 (0.1, 0.57)	0.004
NN50	1.74 (−2.31, 5.8)	3.97	0.56 (0.29, 0.76)	<0.0001	−3.0 (−6.97, 0.97)	3.84	0.48 (0.24, 0.7)	<0.0001
Ln (NN50)	1.26 (0.66, 2.39)	1.76	SV	SV	0.67 (0.32, 1.41)	1.95	SV	SV
pNN50	0.74 (−0.62, 2.10)	1.33	0.54 (0.30, 0.75)	<0.0001	−0.64 (−1.84, 0.57)	1.80	0.50 (0.21, 0.72)	<0.0001
Lomb LF/HF ratio	1.39 (0.58, 2.19)	0.79	0.31 (0.03, 0.58)	0.01	0.74 (0.25, 1.21)	0.44	0.36 (0.11, 0.60)	0.003
Lomb LF (np)	10.12 (2.51, 7.74)	7.45	0.36 (0.08, 0.61)	0.003	9.38 (3.53, 15.22)	5.64	0.38 (0.12, 0.64)	0.001
Ln Lomb LF (np)	1.24 (1.00, 1.54)	0.63	SV	SV	1.26 (1.08, 1.46)	0.51	SV	SV
Lomb HF (np)	−10.12 (−17.74, −2.51)	7.45	0.36 (0.08, 0.61)	0.003	−9.38 (−15.22, −3.53)	5.64	0.38 (0.12, 0.61)	0.001
Ln Lomb HF (np)	0.72 (0.59, 0.89)	0.61	SV	SV	0.79 (0.68, 0.91)	0.50	SV	SV
Poincare SD1	−0.16 (−7.17, 6.86)	6.85	0.39 (0.13, 0.59)	0.001	−0.55 (−4.95, 3.84)	4.17	0.35 (0.09, 0.57)	0.005
Ln Poincare SD1	1.02 (0.81, 1.27)	0.65	SV	SV	1.01 (0.77, 1.32)	0.72	SV	SV
Poincare SD2	3.78 (−8.18, 15.75)	11.69	0.25 (−0.01, 0.5)	0.04	3.28 (−3.09, 9.66)	6.14	0.29 (0.05, 0.51)	0.02
Ln Poincare SD2	1.16 (0.98, 1.36)	0.55	SV	SV	1.10 (0.96, 1.26)	0.49	SV	SV
DFA alpha	0.03 (−0.04, 0.09)	0.06	0.23 (−0.09, 0.48)	0.07	0.03 (−0.03, 0.09)	0.06	0.33 (0.06, 0.58)	0.007
DFA alpha1	0.1 (−0.0004, 0.19)	0.09	0.36 (0.11, 0.6)	0.003	0.11 (0.02, 0.19)	0.08	0.35 (0.09, 0.60)	0.005
DFA alpha2	−0.02 (−0.12, 0.08)	0.09	0.09 (−0.11, 0.3)	0.45	0.02 (−0.06, 0.09)	0.07	0.38 (0.16, 0.58)	0.002
Fractal Dimension	0.01 (−0.05, 0.08)	0.06	−0.02 (−0.3, 0.22)	0.84	0.02 (−0.04, 0.08)	0.06	0.05 (−0.02, 0.27)	0.68
LLE	−0.04 (−0.13, 0.05)	0.08	0.09 (−0.16, 0.33)	0.47	0.08 (−0.03, 0.2)	0.11	0.17 (−0.11, 0.41)	0.17
LZ	−0.04 (−0.09, 0.02)	0.05	0.27 (−0.05, 0.53)	0.03	−0.04 (−0.1, 0.008)	0.05	0.24 (−0.08, 0.52)	0.06
Multiscale Entropy	−0.0002 (−0.32, 0.32)	0.31	0.26 (0.01, 0.49)	0.04	0.01 (−0.29, 0.31)	0.29	0.32 (0.15, 0.48)	0.01
Sample Entropy	−0.07 (−0.19, 0.05)	0.12	0.13 (−0.16, 0.37)	0.29	−0.07 (−0.19, 0.034)	0.11	0.37 (0.15, 0.55)	0.003
Spectral Entropy	−0.15 (−0.32, 0.01)	0.16	0.31 (0.02, 0.57)	0.01	−0.005 (−0.21, 0.20)	0.20	0.20 (−0.07, 0.46)	0.104
Spectral Kurtosis	−0.08 (−2.89, 2.74)	2.71	0.28 (−0.01, 0.53)	0.03	−1.27 (−4.24, 1.71)	2.86	0.30 (−0.01, 0.6)	0.014
Ln Spectral Kurtosis	1.10 (0.85, 1.43)	0.71	SV	SV	0.99 (0.74, 1.31)	0.76	SV	SV

	Reliability between Easy trial Visit 1 and Visit 2				Reliability between Hard trial, Visit 1 and 2			
	Difference of Means (95% CI)	**CR**	**Kendal's tau (95% CI)**	**p-value**	**Difference of Means (95% CI)**	**CR**	**Kendal's tau (95% CI)**	**p-value**
Spectral Skewness	0.132 (−0.29, 0.55)	0.40	0.28 (−0.01, 0.54)	0.02	−0.04 (−0.53, 0.44)	0.47	0.28 (−0.002, 0.57)	0.02

Abbreviations: CI- Confidence Interval; CR- Coefficient of repeatability; Ln- logarithm of the measure; np- normalized percentage; DFA- detrended fluctuation analysis; LLE-largest Lyapunov Exponent; LZ - Lempel-Ziv complexity measure; SV- similar values of Kendall's τ and *p*-values were observed for the log transformed HRV measure. Note: Lower the CR, better the reliability and lower the mean difference, higher the reliability

Table 2

Sensitivity and effect size of HRV parameters to mental task

	Easy trial Visit 1	Hard trial Visit 1	Wilcoxon's Signed rank	Effect size estimates	Easy trial Visit 2	Hard trial Visit 2	Wilcoxon's Signed rank	Effect size estimates
	Mean (SD)	Mean (SD)	(p values)	Gamma	Mean (SD)	Mean (SD)	(p values)	Gamma
Mean RR	0.95 (0.15)	0.92 (0.15)	<0.0001	0.91	0.90 (0.10)	0.88 (0.10)	<0.0001	0.89
SDNN	0.04 (0.03)	0.03 (0.03)	<0.0001	0.48	0.03 (0.02)	0.02 (0.01)	<0.0001	0.64
RMSSD	0.04 (0.04)	0.03 (0.04)	<0.0001	0.80	0.03 (0.03)	0.02 (0.02)	<0.0001	0.71
NN50	17.23	11.05	0.002	0.81	13.37	11.94	0.46	0.89
	(30.44)	(23.38)			(24.62)	(21.73)		
pNN50	5.62 (9.67)	3.72 (8.18)	0.001	0.81	4.24 (7.64)	3.72 (7.10)	0.19	0.83
Lomb LF/HF ratio	2.76 (2.32)	2.12 (1.71)	0.04	0.51	1.65 (1.22)	1.58 (1.48)	0.79	0.53
Lomb LF Power (np)	62.04 (22.73)	59.08 (20.38)	0.91	0.51	53.14 (21.76)	50.53 (21.39)	0.17	0.58
Lomb HF Power (np)	37.96 (22.73)	40.92 (20.38)	0.91	0.51	46.86 (21.76)	49.47 (21.39)	0.17	0.58
Poincare SD1	25.82 (31.63)	21.06 (28.18)	<0.0001	0.80	18.43 (18.61)	14.80 (12.04)	<0.0001	0.71
Poincare SD2	58.19 (42.35)	48.99 (37.50)	<0.0001	0.48	43.48 (33.65)	34.77 (15.91)	<.00001	0.63
DFA alpha	0.94 (0.21)	0.95 (0.19)	0.62	0.35	0.95 (0.17)	0.94 (0.20)	0.66	0.59
DFA alpha1	1.09 (0.34)	1.07 (0.30)	0.34	0.49	1.04 (0.29)	1.01 (0.32)	0.25	0.58
DFA alpha2	0.86 (0.22)	0.93 (0.24)	0.04	0.40	0.90 (0.21)	0.94 (0.20)	0.31	0.30
Fractal Dimension	0.90 (0.07)	0.88 (0.12)	0.90	0.42	0.88 (0.15)	0.86 (0.14)	0.65	**0.07** ϕ
LLE	0.23 (0.12)	0.29 (0.28)	0.48	**−0.03** ϕ	0.27 (0.23)	0.22 (0.12)	0.44	−0.18
LZ	1.29 (0.14)	1.27 (0.17)	0.55	0.33	1.30 (0.13)	1.3 (0.13)	0.70	0.49
Multiscale Entropy	4.54 (0.66)	4.65 (0.71)	0.23	0.33	4.47 (0.91)	4.58 (0.82)	0.32	0.48
Sample Entropy	1.40 (0.26)	1.45 (0.26)	0.25	0.40	1.43 (0.32)	1.49 (0.35)	0.31	0.54
Spectral Entropy	1.89 (0.51)	1.85 (0.45)	0.35	0.42	1.95 (0.47)	1.78 (0.56)	0.03	0.47
Spectral Kurtosis	7.60 (5.24)	8.15 (5.17)	0.26	0.32	8.49 (9.11)	9.99 (9.80)	0.05	0.51
Spectral Skewness	2.00 (1.10)	2.09 (1.07)	0.41	0.37	2.09 (1.20)	2.30 (1.40)	0.18	0.50

Abbreviations: np- normalized percentage; DFA- detrended fluctuation analysis; LLE-largest Lyapunov Exponent; LZ - Lempel-Ziv complexity measure. Note: All the HRV parameters had significant effect sizes with $p \leq 0.005$, except those in bold and marked with **ϕ**.