Interindividual variability in HRV reactivity: autonomic flexibility or regression to the mean?

Dimitriy A. Dimitriev[1](ORCID 0000-0002-8102-7074), Elena V. Saperova[1], Olga S. Indeykina[1], Aleksey D. Dimitriev[1]

[1] Department of Biology, Chuvash State Pedagogical University named I Ya Yakovlev, st. K. Marx, 38, Cheboksary, Chuvash Republic, 428000.

E-mail: rothman68@mail.ru

Abstract

Objective: There is now a substantial body of evidence linking the baseline level of heart rate variability (HRV) with the magnitude of stress-induced reduction in respiratory sinus arrhythmia. However, it remains to be proved to what extent these interindividual differences in stress responses may be attributed to the statistical phenomenon of regression to the mean (RTM). We sought to test the hypothesis that the statistical artifact RTM explains part of the baseline effect. *Approach:* Heart rate recording was carried out in 1,156 volunteers. To obtain an estimate of the stress response, 148 persons were randomly selected. Participants were monitored on a rest day and just before an academic examination for state anxiety and HRV. Participants were divided into quartiles according to baseline HRV levels and were compared for response to academic stress. *Main results:* We observed a significant reduction in HRV in subjects with a high baseline HRV (> 75th percentile), while a significant increase was found in the group with low baseline HRV. Regression analysis demonstrated that the value of baseline HRV correlated with the magnitude of stress reaction consistent with the RTM model. Baseline-adjusted ANCOVA does not reveal significant intergroup differences in the changes in heart rate (HR) and HRV from rest to exam. RTM-adjusted estimates confirmed an exam effect for HR and HRV. *Significance:* The results of our study strongly support RTM as the source of variability of stress-related changes in HRV.

Keywords: regression to the mean, stress, stress reactivity, heart rate variability

1. Introduction

Although the precise nature of emotion is not yet clear, theorists and researchers agree that emotion involves changes across multiple response systems (Gross et al. 2006). According to Cohen et al. (1995), psychological stress occurs when an individual perceives that the situation taxes or exceeds his or her adaptive capacity to cope with it. The model of neurovisceral integration has been widely accepted as a heuristically useful framework for the study of psychological stress (Thayer and Lane 2000). This model conceptualizes that autonomic flexibility provides the basis for adaptation to environmental challenge. Autonomic flexibility can be described as the capacity of the autonomic nervous system (ANS) to adapt to changes in the environment by modifying arousal, respiration, heart rate (HR), and attention (Friedman and Thayer 1998). Porges' proposed polyvagal theory (Porges 1995) postulates that smart vagus (nucleus ambiguous branch of the parasympathetic nervous system) provides the neurophysiological substrate for social communication, attention, and self-regulation (Sinnreich et al. 1998).

Respiratory sinus arrhythmia (RSA), the rhythmical fluctuation in RR intervals accompanying breathing, has been widely used as an index of cardiac vagal tone (Berntson et al. 1993). Low rest RSA and redundant RSA withdrawal in response to emotional challenge are associated with psychopathology, anxiety, and panic disorder (Beauchaine 2001; Licht et al. 2009; Chalmers et al. 2014). These data suggest that RSA can function as a reliable indicator of emotional reactivity. Diminished autonomic flexibility refers to decrease of variability in HR and blunted RSA (Friedman and Thayer 1998). The direct manifestation of flexibility is reactivity of HR and RSA; HR and RSA reactivity can be determined as the difference between stress and rest scores. Thus, a high level of rest RSA should be associated with high RSA reactivity and baseline RSA is a predictor of emotional performance. Observations of stress-induced changes in anxiety and RSA, however, are equivocal.

Kok et al. (2010) found that vagal tone and psychosocial well-being reciprocally and prospectively predict one another. High baseline RSA was related to low trait anxiety (Miu et al. 2009) and high RSA reactivity (Rottenberg et al. 2007). In contrast, Alkozei et al. (2015) reported the absence of differences at rest or in response to stress between anxious and non-anxious children. Gerra et al. (2000) found that adolescent boys with an anxiety disorder showed a higher HR reactivity to a mixed-model stress task than controls. Thus, a higher level of autonomic flexibility was typical for anxious boys. Moreover, there are a growing number of observations that do not fit the theories of Thayer and Porges (Kossowsky et al. 2012; Gorka et al. 2013; Fortunato et al. 2013; Wang et al. 2013; Kristensen et al. 2014; Spangler et al. 2015; Couyoumdjian et al. 2016; Fung et al. 2017). One possible explanation for this disagreement could be the influence of individual differences in autonomic control and autonomic flexibility on RSA reactivity (Berntson and Cacioppo 2004).

The results of the studies outlined above have indicated that autonomic flexibility (i.e., reactivity) depends on the baseline level of RSA (Friedman and Thayer 1998; Beauchaine 2001; Rottenberg et al. 2007; Couyoumdjian et al. 2016). High baseline RSA was associated with a decline in RSA and subjects with low baseline RSA had the opposite reaction to stress. A plausible explanation of this phenomenon is

2

regression to the mean (RTM). There is a strong theoretical rationale for this proposal. The RTM prediction is that extreme baseline levels tend to become less extreme after repeated measurements, even in the absence of real change. Previous studies have demonstrated that RTM is the consequence of random fluctuation or nonsystematic error in repeated measurement (Barnett et al. 2005). Davis (1976) has characterized RTM as the phenomenon by which "a variable that is extreme on its first measurement will tend to be close to the center of the distribution for a later measurement."

The theoretical aspects of RTM are outlined by Lin and Hughes (1997). RTM can be identified by considering the regression (Barnett et al. 2005)

$$\left(Follow-up \ - \ Baseline\right)_i = \beta_0 + \beta_1 \times Baseline + \varepsilon_i .$$

A negative regression coefficient β_1 implies that subjects with very low baseline levels tend to increase and subjects with unusually high baseline levels tend to decrease. This relationship illustrates the RTM effect. Several methods have been proposed for adjusting for the effect of RTM when the variable of interest is supposed to be normally distributed (Lin and Hughes 1997).

Notwithstanding the importance of this issue, the empirical evidence on the effects of RTM on physiological measures is rather sparse and, in particular, the evidence concerning the effects of RTM on autonomic reactivity to psychosocial stress. Some authors consider RTM as a source of bias; they discuss results regarding RTM. Unfortunately, these authors did not test the hypothesis that RTM explains, at least in part, observed changes. One exception is the research of Gotfredsen et al. (1997), who tested for the RTM effect by using a modification of the method of Mee and Chua.

This study aims to evaluate the contribution of RTM and to examine implications for interpreting findings from follow-up HRV research.

2. Methods

2.1 Participants

Data for this study were obtained from research on a cohort conducted from 2005 to 2015. A total of 1,156 students (286 men/870 women, age range from 19 to 24 years (mean age (mean ± SE): 20.53 ± 0.11)) of Chuvash State Pedagogical University participated in the study of heart rate variability (HRV). Each volunteer underwent a comprehensive screening procedure that included physical examination, health history questionnaire, routine laboratory tests, lung function testing, electrocardiogram, and a chest X-ray before the study. There was no evidence of heart or pulmonary disease in any of the subjects. On the day of the study, the volunteers were instructed to avoid alcohol and caffeinated beverages for the 12 preceding hours and to refrain from heavy physical activity since the day before. The local Ethical Committee of biomedical research in Chuvash State University named I. N. Ulyanov approved this study. Informed consent was obtained from all individual participants included in the study.

2.2. State anxiety and heart rate variability

All subjects were administered psychophysiological evaluation composed of baseline (rest) and stressor (just before the academic exam) phases. State anxiety (SA) was assessed using State-Trait Anxiety Inventory anxiety scales (Hanin 1976).

Electrocardiogram (ECG) was recorded continuously throughout each experiment via a three-lead ECG at a 1000 Hz sampling rate. Time series of the interbeat intervals RR were extracted automatically from the ECG signal and analyzed using the Heart Rate Variability Analysis Software (HRVAS) (Ramshur 2010).

Time-domain and frequency-domain measures were derived from RR series, including meanRR; the standard deviation of the normal-to-normal (NN) interval series, SDNN; the spectral power of low-frequency (LF: 0.04–0.15 Hz) and high-frequency (HF: 0.15–0.4 Hz) band power and LF/HF ratio. The LF/HF power ratio, although the subject of discussion as to whether it may be used as an index of the sympatho-vagal interaction (Billman 2013), has been widely used in the study of stress (Castaldo et al. 2015; Hamilton and Alloy 2016).

2.3 Stress due to university examination.

A study of stress reactivity (defined as the difference between exam and rest values) was conducted on a sample of students randomly selected from the original 1,156 subjects. Measures of HRV were natural-log transformed (ln) to meet assumptions of parametric statistics. The sample size was calculated using MedCalc Statistical Software version 17.6 (MedCalc Software bvba, Ostend, Belgium; http://www.medcalc.org; 2017). An admission was made that the study should detect a minimal effect size at a significance level of 0.05 and 80% power (Machin et al. 2011). The sample size was calculated on expected differences (δ) and standard deviations of difference (SD) of HR (δ = 3 beats/min, SD = 11.77 beats/min), lnSDNN (δ = 0.09 s, SD = 0.39 s) , lnLF (δ = 0.25 ms2, SD = 0.95 ms2), lnHF (δ = 0.25 ms2, SD = 1.02 ms2) and lnLF/HF (δ = 0.2, SD = 0.84), on the basis of previous studies from our group (Dimitriev et al. 2008). The minimum sample sizes were estimated to be 123, 148, 116, 133 and 144 for lnHR, lnSDNN, lnLF, lnHF and lnLF/HF, respectively. Therefore, the acceptable sample size was determined to be N 148 for the paired t.

For each variable, the students were divided into three groups, according to the values of the first measurement. The first group of subjects was selected using a baseline measurement of less than the population 25th percentile; the second group represents the middle portion of the distribution, where scores ranged from the 25th percentile to the 75th percentile; and the third group consisted of subjects with baseline measurements greater than the population 75th percentile. Fifteen groups were formed on the basis of HR and HRV, and three groups were formed on the basis of state anxiety. For each variable, students were divided into three groups according to the values of the first measurement: three groups on the basis of HR, three groups on the basis of lnSDNN, three groups on the basis of lnHF, etc. In this way, the same subject could be in different groups depending on the baseline levels of HR and HRV. We analyzed the reaction to stress separately for each group.

4

2.4 Regression to the mean

Ordinary linear-regression analysis was used to identify the RTM. The model was stated as (Follow-up ⬚ Baseline) = constant + $b \times$ baseline. The RTM effect manifests as a negative correlation between baseline values and stress-induced changes in HR and HRV.

Several studies have shown a strong negative correlation between HR and various indices of HRV (time domain and frequency domain) (Billman et al. 2015) that could affect the interpretation of changes in HRV. To correct the effect of heart rate on HRV indices, we used formulae from our previous study (Dimitriev et al. 2016). Corrected lnLF/HF ratio was calculated from corrected lnLF and corrected lnHF. Then linear-regression analysis was performed again to evaluate the effect of baseline HR and HRV on stress reactivity.

The most widely used model of RTM presupposes that the data are normally distributed. Normality of data distribution was tested with the Shapiro–Wilk W test. A test for normality indicated that the HRV variables (SDNN, LF, HF, LF/HF) required log-transformation to meet the criteria for inclusion in the analysis of RTM.

Estimation of RTM assumes that pretest population means μ and variances (σ^2) are known. Population means and variances were calculated from the general baseline data.

Mee and Chua (1991) proposed a modified paired *t*-test that provides a computationally simple approach to separate RTM and stress effects. This approach involves regression of follow-up measurements Y_2 to (baseline level $Y_1 - \mu$):

$$Y_2 = \mu + \tau + \rho(Y_1 - \mu) + \varepsilon,$$

where τ is the intervention effect, ε is the normally distributed random error, and $\rho = \text{Corr}(Y_1; Y_2)$. The null hypothesis of no effect is stated as

$$H_0 : \tau = 0$$

This hypothesis was tested by linear regression analysis of $Y_2 - \mu$ on $Y_1 - \mu$. In this model, the *t*-statistic for testing $\tau = 0$ is the *t*-statistic for testing the hypothesis that the intercept equals 0.

2.5 Statistical analysis

Repeated-measures analysis of variance (ANOVA) with Bonferroni multiple-comparison test was used to examine the effects of the stress on the HR and HRV measures.

Mean values of change between baseline and follow-up were tested for group differences using univariate analysis of covariance (ANCOVA) with "condition" (< 25th percentile, 25th–75th percentiles, > 75th percentile) as between-subject factors and values of the rest period as covariates.

Means of within-subject distributions were tested against the rest values using a *t*-test for a single sample. All data are presented as mean ± SE. Significance was accepted at $p < 0.05$.

3. Results

3.1 Stress due to academic examination

5

The population means of HR, lnSDNN, lnLF, lnHF, and lnLF/HF were 73.11 ± 0.13 bpm, 3.99 ± 0.012 ms, 6.70 ± 0.03 ms^2, 6.81 ± 0.03ms^2, and -0.11 ± 0.02, respectively. Table 1 presents the effect of academic stress on HRV measures in three groups.

Table 1 Comparison of HRV measures and SA between rest and exam stages

Variable	All subjects		1st group (< 25th percentile)		2nd group (25th–75th percentile)		3rd group (> 75th percentile)	
	Rest	Exam	Rest	Exam	Rest	Exam	Rest	Exam
HR [bpm]	73.15 ± 0.89	83.49± 0.98#	60.52 ± 0.60	77.52 ± 2.08#	72.33 ± 0.49	81.59 ± 1.11#	86.96 ± 1.19	92.77 ± 1.57*
lnSDNN [ms]	3.93 ± 0.03	3.78 ± 0.03#	3.49 ± 0.03	3.60 ± 0.05	4.03 ± 0.02	3.77 ± 0.04#	4.52 ± 0.04	4.13 ± 0.07#
lnLF [ms^2]	6.37 ± 0.07	6.21 ± 0.07	5.54 ± 0.07	5.93 ± 0.12#	6.69 ± 0.041	6.26 ± 0.09#	7.73 ± 0.06	6.94 ± 0.16#
lnHF [ms^2]	6.68 ± 0.09	6.03 ± 0.10#	5.38 ± 0.07	5.41 ± 0.18	6.81 ± 0.05	6.07 ± 0.1#	8.15 ± 0.08	6.78 ± 0.26#
lnLF/HF	-0.31 ± 0.06	0.19 ± 0.06#	-1.17 ± 0.05	-0.23 ± 0.10#	-0.09 ± 0.04	0.36 ± 0.08#	0.79 ±0.07	0.55 ± 0.14
SA	25.28 ± 0.82	41.29 ± 0.83#	13.48 ± 0.75	36.76 ± 1.66#	24.30 ± 0.49	40.86 ± 1.2#	39.39 ± 0.75	46.33 ± 1.66#

Values are means ± SE.

Exam vs. rest: $*p < 0.05$; $\#p < 0.01$

The HF component of HRV was significantly reduced in the exam session compared with the rest session ($p < 0.001$), and the HR ratio was significantly higher in the exam session compared with the rest session ($p < 0.001$) (see Table 1). There was no difference in the LF component of HRV between the exam session and the rest session. lnSDNN reduced significantly from baseline during the exam session ($p < 0.001$).

Academic stress resulted in a significant increase in HR in all three groups. The lnSDNN decreased significantly before the examination in the second and third groups, but not in the group with a low baseline level ($p > 0.05$). The passage from the rest session to the exam evoked a significant increase of lnLF in the first group ($p < 0.01$), whereas lnLF in the other two groups decreased significantly ($F = 23.46$, $p < 0.01$).

Statistical analysis of lnHF revealed a significant decrease in this parameter in the second and third groups and insignificant changes of LnHF were observed in the first group ($F = 15.39$, $p < 0.01$). Analysis of SA changes during the exam session yielded an overall increase in SA. It is clear from Table 1 that the first group has a different pattern of HRV changes.

Results of the population HRV analysis indicate that women had on average a higher heart rate than men (73.94 ± 0.36 vs. 70.59 ± 0.63 bmp, $p < 0.01$) and lower lnSDNN (3.97 ± 0.01 vs 4.06 ± 0.02 ms, $p < 0.01$). However, lnSDNN differences disappear after heart rate correction of HRV. Compared to men, women's HR corrected lnSDNN was 3.11 ± 0.34 versus 3.12 ± 0.34 ms ($p > 0.05$). LF and HF did not differ significantly by gender ($p > 0.05$). HRV reactivity to stress was not significantly different between

men and women, and stress-related reduction of lnSDNN was similar in both gender groups (men: -0.22 ± 0.11; women -0,15 ± 0.04 ms; p > 0.05). Comparing stress-related decrease in lnHF between women (-0.66 ± 0.11 ms^2) and men (-0.63 ± 0.30 ms^2), we found no evidence of a gender difference in stress reactivity (p > 0.05). A similar result was obtained for stress-related reduction of lnLF (men: -0.23 ± 0.2; women: -0.15 ± 0.1 ms^2, p > 0.05).

3.2 Regression to the mean

To reveal potentially significant associations between baseline values and stress effect, we have examined scatter plots of change (exam minus rest measurements) against rest measurement (Barnett et al. 2005) (Fig. 1).
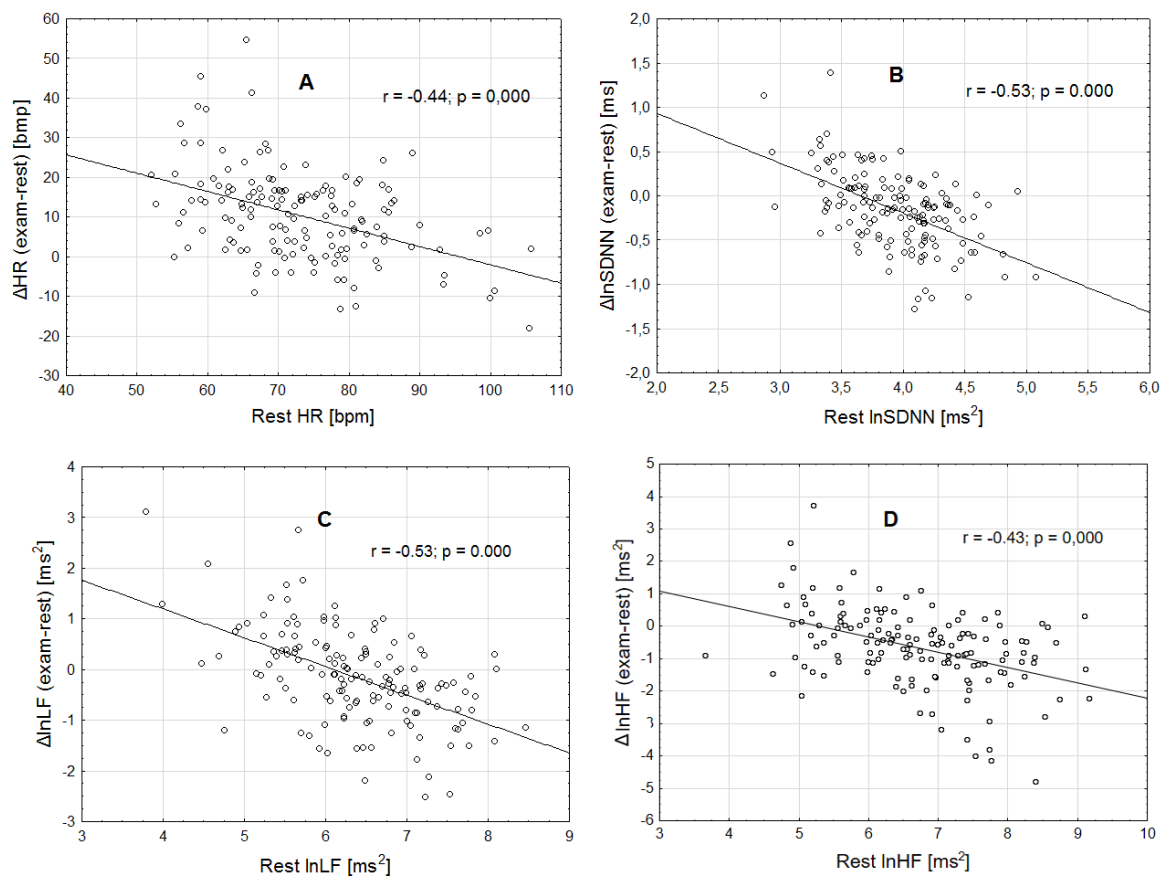

Figure 1 Change (exam – rest) is highly correlated to the baseline level. Scatter plots of baseline HR (A), lnSDNN (B), lnLF (C), and lnHF (D) vs. change. Linear-regression analysis produced significant negative r values for each metric.

Figure 1 shows the changes in HRV measures plotted against baseline levels. For participants, the amount of HRV reduction increased significantly as baseline lnSDNN, lnLF, and lnHF increased (*p* < 0.001 for linear trend for all measures). RTM is evident on the plot, as subjects from the first group have

tended to increase HRV, and subjects with very high baseline HRV have tended to decrease lnSDNN, lnLF, and lnHF. The results for baseline and change ln LF/HF display a similar regression tendency ($r=-0.56$, $b=-0.58$, $p < 0.001$).

The changes in HRV, after correction for heart rate, correlated significantly with baseline level (Table 2).

Table 2 Associations between baseline and changes in heart rate corrected HRV.

Variable	Adjusted R^2	β for baseline	t	P
ln SDNN	0.37	-0.64	9.27	<0.001
ln LF	0.33	-0.61	8.62	<0.001
ln HF	0.22	-0.54	6.51	<0.001
ln LF/HF	0.29	-0.61	7.83	<0.001

To test whether the HRV changes are due to RTM, we applied Mee and Chua's test. Mee–Chua tests for the HR, lnSDNN, and lnHF confirmed an exam effect ($t = 12.11$, $p < 0.001$, $t = 6.40$, $p < 0.001$, $t = 7.88$, $p < 0.001$, respectively). We find a significant difference in lnLF between the exam and rest session after adjusting for RTM (Mee–Chua $t = 6.4$, $p < 0.001$).

For HR and HRV, the rest level was used as a covariate in an ANCOVA model to estimate the effect of RTM. Table 3 presents the descriptive statistics of change between sessions and the results of the ANCOVA.

Table 3 HRV changes between rest and exam (exam – rest) using the ANOVA and ANCOVA

Variable	*Group*	*ANOVA*		*ANCOVA*	
		Mean changes (±SE)	*p**	**Mean changes (±SE)**	*p**
HR [bpm]	1st group (< 25th percentile)	17.00 ± 2.05	< 0.001	8.82 ± 2.71	> 0.05
	2nd group (25th–75th percentile)	9.26 ± 1.19		8.66 ± 1.24	
	3rd group (> 75th percentile)	5.81 ± 1.63		14.60 ± 2.83	
lnSDNN [ms]	1st group (< 25th percentile)	0.11 ± 0.06	< 0.001	-0.27 ± 0.10	> 0.05
	2nd group (25th–75th percentile)	-0.26 ± 0.04		-0.25 ± 0.04	
	3rd group (> 75th percentile)	-0.39 ± 0.07		-0.02 ± 0.11	
lnLF [ms^2]	1st group (< 25th percentile)	0.39 ± 0.12	< 0.001	-0.16 ± 0.20	> 0.05
	2nd group (25th –75th percentile)	-0.43 ± 0.09		-0.41 ± 0.09	
	3rd group (> 75th percentile)	-0.79 ± 0.15		-0.26 ± 0.24	
lnHF [ms^2]	1st group (< 25th percentile)	0.027 ± 0.17	< 0.001	-0.39 ± 0.33	> 0.05
	2nd group (25th–75th percentile)	-0.74 ± 0.11		-0.73 ± 0.13	
	3rd group (> 75th percentile)	-1.37 ± 0.24		-0.97 ± 0.34	

lnLF/HF	1st group (< 25th percentile)	0.94 ± 0.10	< 0.001	0.38 ± 0.17	> 0.05
	2nd group (25th–75th percentile)	0.45 ± 0.09		0.59 ± 0.09	
	3rd group (> 75th percentile)	-0.24 ± 0.14		0.48 ± 0.23	

Values are means ± SE.

*effect of group

When ANCOVA was performed to evaluate the effects of group on changes in HR and HRV, no significant effects were found regarding the changes of HR, lnSDNN, lnLF, and lnHF scores from rest to exam (HR: $F = 2.03$, $p = 0.13$; lnSDNN: $F = 2.26$, $p = 0.11$; lnLF: $F = 1.14$, $p = 0.32$; lnHF: $F = 0.49$, $p = 0.61$, lnLF/HF: $F = 0.98$, $p = 0.37$) after adjusting for the baseline level (Table 3).

The intraindividual distributions of lnHF (Fig. 2) are given below to illustrate the nature of RTM (data from our previous study (Dimitriev DA et al. 2007)). These subjects took part in a study of the relationships between phases of the menstrual cycle and HRV. RR intervals were recorded daily throughout the one month (excluding weekends). The same subjects were examined for acute stress reaction.
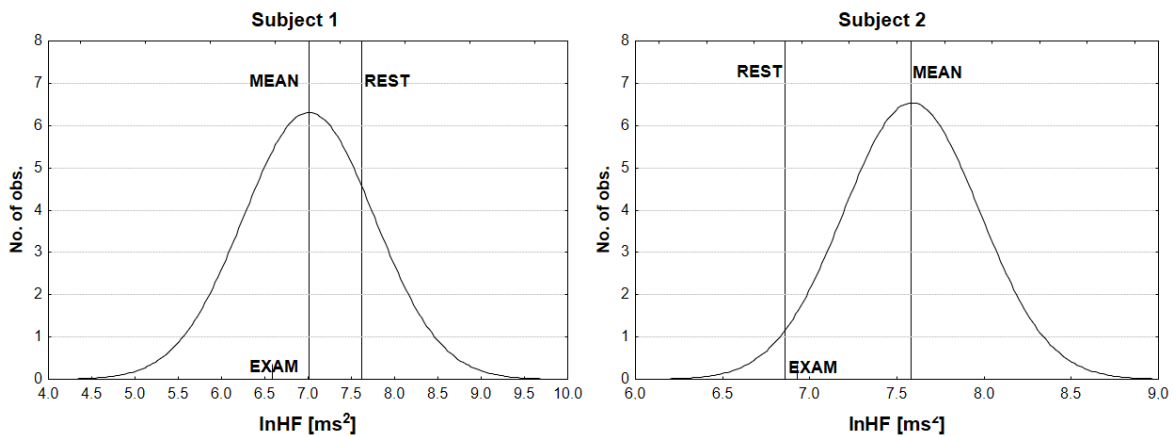

Figure 2 Within-subject distributions and RTM. A mean was calculated from daily lnHF values.

Subject 1's lnHF at rest (7.62 ms^2) was significantly higher than the true mean (7.01 ms^2; $t = 7.18$; $p < 0.001$). The difference between rest and exam was 1.03 ms^2 and between mean and exam 0.42 ms^2. This suggests the stress reaction is overestimated by about 0.61 ms^2. For the second subject, lnHF at rest (6.86 ms^2) was significantly lower than the true mean (7.58 ms^2, $t = 10.51$, $p < 0.001$) and was approximately equal to the lnHF before the exam (6.93 ms^2). We might think that the subject's lnHF at exam had slightly increased when in fact the rest measurement was just unusually low, and the subject's true response was negative (lnHFexam – lnHFmean = –0.65 ms^2; $t = 9.49$ $p < 0.01$). SA was increased in both subjects (Subject 1: from 28 to 42, Subject 2: from 30 to 38).

4. Discussion

A common study design in physiological research is to give subjects some intervention and then observe the response. For example, HRV may be measured twice: once at rest and once during stress. Despite its

evident simplicity, the analysis of this form of study may produce biased results due to RTM, although this is not widely appreciated in the literature and has not been quantified. Stress studies in humans reported that stressful life events are associated with a more prominent depression of RSA in individuals with higher baseline HRV, but to our knowledge, none of these analyses included correction for the RTM effect.

A linear regression analysis was used to assess the relation of stress response to baseline HRV value. The results of this analysis show the difference between rest and exam was influenced by baseline level of HRV: subjects with unusually high baseline levels have more prominent stress response than subjects with very low initial values of HRV variables. These findings indicate the strong RTM effect.

Mee–Chua tests for HRV variables confirmed the marked decrease in SDNN and HF. The crude difference in LF between rest and exam almost reached the significance level ($p = 0.05$) and p-value drawn from the Mee–Chua test was below 0.025. The authors attribute the observed difference to the RTM effect, leading to LF increasing in the first group, which includes subjects with low baseline levels of LF.

Two analytical strategies have emerged to examine changes between baseline and follow-up measurements of HRV. The first is one-way repeated measures analysis of variance (repeated-measures ANOVA) and the second adjusted for baseline HRV data (ANCOVA). ANCOVA enables removal of the RTM effect and permits estimation of adjusted group means. In this study of HRV reactivity, these two approaches lead to different conclusions: the results of the ANOVA suggest a greater decline of HRV in the third group (very high baseline level) and an increase of HRV variables in the first group (very low baseline level). In contrast, those obtained using ANCOVA with a correction to baseline level suggest the similarity of reactions in all three groups. A source of difference between these two approaches is that ANOVA results are biased by RTM.

RTM is omnipresent in human studies (Gavish et al. 2011; Guenancia et al. 2013; McCambridge et al. 2014; Margaritelis et al. 2016; Forstmeier et al. 2016; Perry and Karpova 2017), and the RTM effect can lead the researcher astray. It most commonly occurs in studies where subjects are selected because they expressed very high or very low baseline values. Thus, participants are eligible to be in the study group only if they have unusually extreme values of RSA or some other variable thought to be related to the stress response. By RTM, subjects with an extremely high level of baseline HRV will have a more prominent stress response, regardless of the severity of the state anxiety. In this case, an individual autonomic flexibility is simply owing to a shift in RSA on the second occasion to the individual mean.

Mee and Chua (1991) presented a modified paired t-test that ensures a computationally simple approach to hypothesis testing that accounts for the effect of RTM. The main drawback of this method is that it requires knowledge of the true mean μ and variance σ^2 in the target population. In our research population, mean and variance have been estimated from HRV measurements of 1,153 students, and our number of subjects is comparable to those used in other cohort studies of HRV (Kawachi et al. 1995; Tsuji et al. 1996; Sinnreich et al. 1998; Antelmi et al. 2004). However, such data may not always be available, and it is not common to have population data as a basis for stress research protocols. To overcome

10

this problem, Ostermann et al. (2008) developed a straightforward method based on Mee and Chua's modified *t*-test for the situation, where the population μ is unknown.

RTM is caused by random fluctuations in physiological parameters. Random variation is the substantial feature in HR behavior (Laitio 2007; Castiglioni et al. 2009). The HRV can be affected by various factors, such as physical activity, caffeine beverages, menstrual cycle, mood, alertness, mental activity, respiratory rate, and depth. It is easy to control for recreational activity, natural cycles, caffeine consumption, and respiration, but we cannot completely rule out that the person's emotional and mental state may bias the results of HRV measurements.

The RTM effect can be corrected by adequate study design and/or appropriate adjustment. It has been recognized that multiple baseline measurements could significantly reduce the RTM effect (Barnett et al. 2005). Because of the study design, we could not implement an RTM correction through multiple baseline measurements, which could be regarded as a limitation of our study. We found it to be evident from a wide variety of research papers that the discussion of RTM affects many fields of physiology. Unfortunately, methods to adjust for RTM are rarely used in the analysis of physiological data. We, therefore, believe that the lack of adjustment to RTM in most of the studies probably leads to a significant bias in the results.

5. Conclusions

Studies of autonomic flexibility demonstrated that high baseline HRV was associated with a decline in HRV and subjects with low baseline HRV had an inverse reaction to stress, suggesting that baseline HRV is a marker of stress adaptive capacity. The results of our study strongly support an alternative view: RTM is the major source of variability of stress-related changes in HRV. The effect of baseline HRV level disappears after adjustment for RTM.

References

Alkozei, A., Creswell, C., Cooper, P.J. and Allen, J.J., 2015. Autonomic arousal in childhood anxiety disorders: associations with state anxiety and social anxiety disorder. *J. Affect. Disord.*, *175*, pp.25-33.doi: 10.1016/j.jad.2014.11.056

Antelmi, I., De Paula, R.S., Shinzato, A.R., Peres, C.A., Mansur, A.J. and Grupi, C.J., 2004. 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.*, *93*(3), pp.381-385.doi: 10.1016/j.amjcard.2003.09.065

11

Barnett, A.G., Van Der Pols, J.C. and Dobson, A.J., 2004. Regression to the mean: what it is and how to deal with it. *Int. J. Epidemiol.* ,*34*(1), pp.215-220.

Beauchaine T 2001 2001. Vagal tone, development, and Gray's motivational theory: Toward an integrated model of autonomic nervous system functioning in psychopathology. Dev. Psychopathol. *13*(2), pp.183-214.

Berntson, G.G., Cacioppo, J.T. and Quigley, K.S., 1993. Respiratory sinus arrhythmia: autonomic origins, physiological mechanisms, and psychophysiological implications. *Psychophysiology*, *30*(2), pp.183-196. Berntson, G.G. and Cacioppo, J.T., 2004. Heart rate variability: Stress and psychiatric conditions. In: Malik M, Camm AJ (ed) Dynamic electrocardiography. Futura; New York, pp 57-64

Billman, G.E., 2013. The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. *Frontiers in physiology*, *4*. 26. doi: 10.3389/fphys.2013.00026

Billman, G.E., Huikuri, H.V., Sacha, J. and Trimmel, K., 2015. An introduction to heart rate variability: methodological considerations and clinical applications. *Frontiers in physiology*, *6*.:55. doi: 10.3389/fphys.2015.00055

Castaldo, R., Melillo, P., Bracale, U., Caserta, M., Triassi, M. and Pecchia, L., 2015. Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis. *Biomedical Signal Processing and Control*, *18*, pp.370-377. doi: 10.1016/j.bspc.2015.02.012

Castiglioni, P., Parati, G., Civijian, A., Quintin, L. and Di Rienzo, M., 2009. Local scale exponents of blood pressure and heart rate variability by detrended fluctuation analysis: effects of posture, exercise, and aging. *IEEE Transactions on Biomedical Engineering*, *56*(3), pp.675-684. doi: 10.1109/TBME.2008.2005949

Chalmers, J.A., Quintana, D.S., Maree, J., Abbott, A. and Kemp, A.H., 2014. Anxiety disorders are associated with reduced heart rate variability: a meta-analysis. *Frontiers in psychiatry*, *5*.:80. doi: 10.3389/fpsyt.2014.00080

Cohen, S., Kessler, R.C. and Gordon, L.U., 1995 Strategies for measuring stress in studies of psychiatric and physical disorders. In: Kessler RC, Gordon UL, Cohen S (ed) Measuring stress: A guide for health and social scientists. Oxford University Press, New York, pp 3-26

Couyoumdjian, A., Ottaviani, C., Petrocchi, N., Trincas, R., Tenore, K., Buonanno, C. and Mancini, F., 2016. Reducing the meta-emotional problem decreases physiological fear response during exposure in phobics. *Frontiers in psychology*, *7*. 1105. doi: 10.3389/fpsyg.2016.01105

Davis, C.E., 1976. The effect of regression to the mean in epidemiologic and clinical studies. *American journal of epidemiology*, *104*(5), pp.493-498.

12

Dimitriev, D.A., Dimitriev, A.D., Karpenko, Y.D. and Saperova, E.V., 2008. Influence of examination stress and psychoemotional characteristics on the blood pressure and heart rate regulation in female students. *Human Physiology*, *34*(5), pp.617-624.

Dimitriev, D.A., Saperova, E.V., Dimitriev, A.D. and Karpenko, I., 2007. Features of cardiovascular functioning during different phases of the menstrual cycle. *Rossiiskii fiziologicheskii zhurnal imeni IM Sechenova*, *93*(3), pp.300-305.

Dimitriev, D.A., Saperova, E.V. and Dimitriev, A.D., 2016. State anxiety and nonlinear dynamics of heart rate variability in students. *PloS one*, *11*(1), p.e0146131. doi: 10.1371/journal.pone.0146131

Forstmeier, W., Wagenmakers, E.J. and Parker, T.H., 2016. Detecting and avoiding likely false-positive findings—a practical guide. Biol Rev Camb Philos Soc. doi:10.11111/brv.12315.

Fortunato, C.K., Gatzke-Kopp, L.M. and Ram, N., 2013. Associations between respiratory sinus arrhythmia reactivity and internalizing and externalizing symptoms are emotion specific. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(2), pp.238-251. doi: 10.3758/s13415-012-0136-4

Friedman, B.H. and Thayer, J.F., 1998. Anxiety and autonomic flexibility: a cardiovascular approach. *Biological psychology*, *47*(3), pp.243-263.

Fung, B.J., Crone, D.L., Bode, S. and Murawski, C., 2017. Cardiac signals are independently associated with temporal discounting and time perception. *Frontiers in behavioral neuroscience*, *11*.doi: 10.3389/fnbeh.2017.00001

Gavish, B., Alter, A., Barkai, Y., Rachima-Maoz, C., Peleg, E. and Rosenthal, T., 2011. Effect of non-drug interventions on arterial properties determined from 24-h ambulatory blood pressure measurements. *Hypertension Research*, *34*(11), pp.1233-1238. doi: 10.1038/hr.2011.125

Gerra, G., Zaimovic, A., Zambelli, U., Timpano, M., Reali, N., Bernasconi, S. and Brambilla, F., 2000. Neuroendocrine responses to psychological stress in adolescents with anxiety disorder. *Neuropsychobiology*, *42*(2), pp.82-92.

Gorka, S.M., Nelson, B.D., Sarapas, C., Campbell, M., Lewis, G.F., Bishop, J.R., Porges, S.W. and Shankman, S.A., 2013. Relation between respiratory sinus arrythymia and startle response during predictable and unpredictable threat. *Journal of psychophysiology*. doi: 10.1027/0269-8803/a000091

Gotfredsen, A., Bæksgaard, L. and Hilsted, J., 1997. Body composition analysis by DEXA by using dynamically changing samarium filtration. *Journal of Applied Physiology*, *82*(4), pp.1200-1209.

Gross, J.J., Richards, J.M. and John, O.P., 2006. 2006. Emotion regulation in everyday life. In: Snyder DK, Simpson JE, Hughes JN (eds) Emotion regulation in couples and families: Pathways to dysfunction and health, edited by American Psychological Association; Washington, DC, pp 13-35

13

Predictors of post-stress LVEF drop 6 months after reperfused myocardial infarction: a gated myocardial perfusion SPECT study, doi: 10.1007/s12149-012-0661-9

Hamilton, J.L. and Alloy, L.B., 2016. Atypical reactivity of heart rate variability to stress and depression across development: Systematic review of the literature and directions for future research. *Clinical psychology review*, *50*, pp.67-79. doi: 10.1016/j.cpr.2016.09.003

Hanin Y.L., 1976 Quick Guide to the application of the scale of reactive and personal anxiety BH Spielberger. LNIIFK, Leningrad

Kawachi, I., Sparrow, D., Vokonas, P.S. and Weiss, S.T., 1995. Decreased heart rate variability in men with phobic anxiety (data from the Normative Aging Study). *The American journal of cardiology*, *75*(14), pp.882-885.

Kok, B.E. and Fredrickson, B.L., 2010. Upward spirals of the heart: Autonomic flexibility, as indexed by vagal tone, reciprocally and prospectively predicts positive emotions and social connectedness. *Biological psychology*, *85*(3), pp.432-436. doi: 10.1016/j.biopsycho.2010.09.005

Kossowsky, J., Wilhelm, F.H., Roth, W.T. and Schneider, S., 2012. Separation anxiety disorder in children: disorder specific responses to experimental separation from the mother. *Journal of Child Psychology and Psychiatry*, *53*(2), pp.178-187. doi: 10.1111/j.1469-7610.2011.02465.x

Kristensen, H., Oerbeck, B., Torgersen, H.S., Hansen, B.H. and Wyller, V.B., 2014. Somatic symptoms in children with anxiety disorders: an exploratory cross-sectional study of the relationship between subjective and objective measures. *European child & adolescent psychiatry*, *23*(9), pp.795-803. doi: 10.1007/s00787-013-0512-9

Laitio, T., Jalonen, J., Kuusela, T. and Scheinin, H., 2007. The role of heart rate variability in risk stratification for adverse postoperative cardiac events. *Anesthesia & Analgesia*, *105*(6), pp.1548-1560. doi: 10.1213/01.ane.0000287654.49358.3a

Licht, C.M., De Geus, E.J., Van Dyck, R. and Penninx, B.W., 2009. Association between anxiety disorders and heart rate variability in The Netherlands Study of Depression and Anxiety (NESDA). *Psychosomatic medicine*, *71*(5), pp.508-518. doi: 10.1097/PSY.0b013e3181a292a6

Lin, H.M. and Hughes, M.D., 1997. Adjusting for regression toward the mean when variables are normally distributed. *Statistical Methods in Medical Research*, *6*(2), pp.129-146.

Machin D., Campbell M.J., Tan S.B., Tan S.H., 2011 Sample size tables for clinical studies. John Wiley & Sons, Oxford

Margaritelis, N.V., Theodorou, A.A., Paschalis, V., Veskoukis, A.S., Dipla, K., Zafeiridis, A., Panayiotou, G., Vrabas, I.S., Kyparos, A. and Nikolaidis, M.G., 2016. Experimental verification of re-

gression to the mean in redox biology: differential responses to exercise. *Free radical research*, *50*(11), pp.1237-1244. doi: 10.1080/10715762.2016.1233330

McCambridge, J., Kypri, K. and McElduff, P., 2014. Regression to the mean and alcohol consumption: a cohort study exploring implications for the interpretation of change in control groups in brief intervention trials. *Drug and alcohol dependence*, *135*, pp.156-159. doi: 10.1016/j.drugalcdep.2013.11.017

Mee, R.W. and Chua, T.C., 1991. Regression toward the mean and the paired sample t test. *The American Statistician*, *45*(1), pp.39-42.

Miu, A.C., Heilman, R.M. and Miclea, M., 2009. Reduced heart rate variability and vagal tone in anxiety: trait versus state, and the effects of autogenic training. *Autonomic Neuroscience*, *145*(1), pp.99-103. doi: 10.1016/j.autneu.2008.11.010

Ostermann, T., Willich, S.N. and Lüdtke, R., 2008. Regression toward the mean–a detection method for unknown population mean based on Mee and Chua's algorithm. *BMC medical research methodology*, *8*(1), p.52. doi:10.1186/1471-2288-8-52

Efficacy of Teaching Creative Thinking Skills: A Comparison of Multiple Creativity Assessments doi: 10.1016/j.tsc.2015.169

Porges, S.W., 1995. Orienting in a defensive world: Mammalian modifications of our evolutionary heritage. A polyvagal theory. *Psychophysiology*, *32*(4), pp.301-318.

Ramshur J.T., 2010. Design, evaluation, and application of heart rate variability software (HRVAS). Master's thesis. The University of Memphis

Rottenberg, J., Clift, A., Bolden, S. and Salomon, K., 2007. RSA fluctuation in major depressive disorder. *Psychophysiology*, *44*(3), pp.450-458. doi: 10.1111/j.1469-8986.2007.00509.x

Sinnreich, R., Kark, J.D., Friedlander, Y., Sapoznikov, D. and Luria, M.H., 1998. Five minute recordings of heart rate variability for population studies: repeatability and age–sex characteristics. *Heart*, *80*(2), pp.156-162.

Spangler, D.P., Bell, M.A. and Deater☐Deckard, K., 2015. Emotion suppression moderates the quadratic association between RSA and executive function. *Psychophysiology*, *52*(9), pp.1175-1185. doi: 10.1111/psyp.12451

Thayer, J.F. and Lane, R.D., 2000. A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of affective disorders*, *61*(3), pp.201-216.

Tsuji, H., Larson, M.G., Venditti, F.J., Manders, E.S., Evans, J.C., Feldman, C.L. and Levy, D., 1996. Impact of reduced heart rate variability on risk for cardiac events. *Circulation*, *94*(11), pp.2850-2855.

Wang, Z., Lü, W. and Qin, R., 2013. Respiratory sinus arrhythmia is associated with trait positive affect and positive emotional expressivity. *Biological psychology*, *93*(1), pp.190-196. doi: 10.1016/j.biopsycho.2012.12.006