



Published in final edited form as:

*Biol Psychol.* 2012 February ; 89(2): 349–364. doi:10.1016/j.biopsycho.2011.11.009.

## Statistical strategies to quantify respiratory sinus arrhythmia: Are commonly used metrics equivalent?

Gregory F. Lewis, PhD<sup>1</sup>, Senta A. Furman, BS<sup>1</sup>, Martha F. McCool, MA<sup>1</sup>, and Stephen W. Porges, PhD<sup>1</sup>

<sup>1</sup>Brain-Body Center, Department of Psychiatry, University of Illinois at Chicago, Chicago, Illinois, USA.

### Abstract

Three frequently used RSA metrics are investigated to document violations of assumptions for parametric analyses, moderation by respiration, influences of nonstationarity, and sensitivity to vagal blockade. Although all metrics are highly correlated, new findings illustrate that the metrics are noticeably different on the above dimensions. Only one method conforms to the assumptions for parametric analyses, is not moderated by respiration, is not influenced by nonstationarity, and reliably generates stronger effect sizes. Moreover, this method is also the most sensitive to vagal blockade. Specific features of this method may provide insights into improving the statistical characteristics of other commonly used RSA metrics. These data provide the evidence to question, based on statistical grounds, published reports using particular metrics of RSA.

### Keywords

heart rate variability; respiratory sinus arrhythmia; heart rate; respiration; autonomic nervous system

### Introduction

Respiratory sinus arrhythmia (RSA) is a frequently measured physiological metric studied in several basic science and clinical disciplines. Although the methods used to quantify RSA are not standardized and vary among laboratories, little effort has been directed at establishing appropriate criteria with which to contrast methods. The published “standards” for quantifying heart rate variability (HRV) in cardiology (Camm, Malik, Bigger, Breithardt, Cerutt, Cohen, et. al, 1996) and psychophysiology (Berntson, Bigger, Eckberg, Grossman, Kauffmann, Malik, et. al, 1997) have not prioritized the methods used to quantify RSA on either statistical or neurophysiological bases. In the absence of critical criteria to evaluate RSA metrics, researchers have assumed, based on reports of high within-subject correlations between common RSA metrics that all methods are equivalent (e.g., Grossman, van Beek, & Wientjes, 1990; Goedhart, 2007). However, within subject correlation is a deficient methodology to establish statistical equivalence between measures, since the within-subject

© 2011 Elsevier B.V. All rights reserved.

*Corresponding Author:* Stephen W. Porges, Department of Psychiatry (MC912), 1601 W Taylor Street, Chicago, IL 60608, USA, sporges@uic.edu, phone: 312-355-1557.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

correlation between biased and unbiased metrics can approach unity and the distributional features of each metric will not influence the strength of the linear relationship (Altman & Bland, 1983; Bland & Altman, 1986).

Physiologists have known for a century that vagal cardioinhibitory fibers have a respiratory rhythm (Hering, 1910). Although the functional impact of this vagal rhythm produces RSA, there has been a continuous debate about the efficacy of using RSA as a dynamic index of cardiac vagal tone. Instead there has been a dependence on using a heart rate measure, often in response to blockade or surgical disruption of vagal efferent outflow to the heart, and not RSA as the criterion measure of the functional impact of vagal cardioinhibitory fibers on the heart (Katona & Jih, 1975; Fouad, Tarazi, Ferrario, Fighaly, & Alicandri, 1984; Grossman & Kollai, 1993). During the past 30 years, research has expanded our understanding of the neural circuits that regulate cardioinhibitory vagal pathways. It is now accepted that, in humans and other mammals, the primary cardioinhibitory pathways originate in the nucleus ambiguus (Rentero, et. al, 2002). These vagal pathways are myelinated, have nicotinic preganglionic receptors, and have a respiratory rhythm (see Porges 2007 for a review). Thus, based on current knowledge of neurophysiology, a strong argument can be made that RSA reflects the dynamic functional impact of the vagal fibers originating in the nucleus ambiguus.

Other vagal cardioinhibitory pathways originate in the dorsal motor nucleus of the vagus. These fibers are unmyelinated. Although the output of these fibers does not have a respiratory rhythm, they contribute to the bradycardia associated with hypoxia and baroreceptor reflexes. Less is known about the preganglionic receptors of the unmyelinated vagal efferents that originate from the dorsal motor nucleus of the vagus and influence heart rate. Since the influence of unmyelinated vagal fibers on heart rate is preserved following nicotinic blockade, it has been proposed that the preganglionic receptors for these pathways are muscarinic (Cheng & Powley, 2000). However, this explanation is still questioned and other mechanisms may be involved in the regulation of the preganglionic receptors of the unmyelinated vagus.

Many psychophysiolgists have assumed that the quantitative metrics used by bench physiologists could be generalized and ported to human research. Unfortunately for psychophysiolgists, who are applying these physiological measures in experimental paradigms, the physiologists neither evaluated whether statistical parameters of RSA metrics conformed to the assumptions necessary for parametric analyses nor compared the relative sensitivity and specificity of various metrics to vagal manipulation. In addition, although specific quantification technologies for RSA, such as paced breathing protocols, may provide insight into clinical pathologies for physicians (e.g., Low and Sletten, 2008) and into cardiopulmonary interactions for physiologists (e.g., Hirsch and Bishop, 1981), these technologies may be inappropriate when applied to dynamically moving and psychologically active humans.

The research described in this paper provides an important step in identifying criteria to evaluate the relative merits of methods used to quantify RSA. Analyses will be presented that contrast three common methods used to quantify RSA in psychophysiological research and will describe how the different RSA metrics conform to statistical assumptions, are moderated by respiration, distorted by trend, and are sensitive to vagal blockade. To investigate these differences, RSA metrics were quantified during baseline and during infusion to either saline or a cholinergic blockade. These data enable analyses to address nine specific research questions that may contribute to an understanding of the appropriateness of each RSA metric. First, are the measures correlated? Second, are the methods equivalent in test-retest reliability and can reliable estimates be generated over

short time periods? Third, do distributions of the measures conform to the assumption of normality? Fourth, do the measures differ in their direct relation to respiratory parameters? Fifth, are the methods differentially influenced by violations of the stationarity assumption? Sixth, is there a difference in sensitivity and statistical power of the RSA measures to a peripheral blockade of vagal outflow to the heart? Seventh, as an additional index of vagal modulation, is the moderation by respiratory parameters of the relation between change in RSA and change in heart rate in response to a saline infusion metric dependent? Eighth, is the relationship among the three RSA metrics moderated by respiration or trend? Ninth, are all methods equivalent when appropriately detrended and the estimates of variance are logarithmically transformed? The answers to these questions will provide insight into the disparate findings in the literature and will identify specific quantitative strategies that may improve the psychometric features of the less robust metrics.

## Methods

### Subjects

Sixty-five male participants between the ages of 18 and 34 ( $M = 25.48$ ,  $SD = 3.99$ ) were recruited with flyers, Craigslist advertisements, literature distributed at the UIC Hospital, and the UIC psychology student subject pool. Participants self-identified as Caucasian (58.5%), African-American (21.5%), Asian (10.8%), or other (9.2%) and were excluded from the study if, in the preceding 24 hours, they had used a tobacco product, consumed more than 3 alcoholic beverages, taken any non-prescription drugs, or had a caffeine drink within the two hours prior to the experimental session. Potential participants were interviewed by phone prior to arriving at the research site. Phone screening included questions on health habits and general medical health. Before participating in sessions involving intravenous infusions, a medical screening was conducted that included a 12-lead ECG performed by clinical staff and reviewed by a study physician. No participant had an existing medical condition that would place him at risk or confound the autonomic measurements. In addition, no participant was taking prescription medications, including central nervous system depressants or stimulants, hypertension medications, or anti-cholinergic agents that could influence autonomic regulation.

### Protocol

The research protocol required participation in two test sessions. One session was conducted in a typical psychophysiological research laboratory located in an office building and the other session was conducted in a clinical research center located in a hospital. Participants were randomly assigned to either the hospital or office building for their first session and rescheduled in approximately two weeks for their second session. During the hospital session participants received an intravenous bolus infusion of either saline vehicle or glycopyrrolate (.006mg/kg) following the initial five-minute pre-testing baseline. Glycopyrrolate is a nonspecific peripheral muscarinic cholinergic receptor antagonist. The drug does not cross the blood-brain barrier. All peripheral postganglionic receptors of vagal efferent pathways to cardiac tissue are muscarinic. The recommended effective clinical dose was used.

Respiratory and beat-to-beat heart rate data collected during both the baseline condition at the research laboratory setting and the pre-infusion and post-infusion glycopyrrolate conditions at the hospital setting were analyzed to evaluate and to contrast the distributional characteristics and sensitivity of three common metrics of RSA to respiratory parameters, trend, and vagal blockade. Data were collected from 65 participants. 47 participants had complete data from two seated baseline sessions in both the laboratory and hospital sessions. Sample sizes were maximized in the analyses presented below (i.e., in the research

laboratory 65 participants were tested at the first baseline and 48 participants at the second baseline, in the hospital 50 participants were tested). Of the 50 participants tested in the hospital, 25 received glycopyrrolate infusion and 25 received a “control” saline infusion.

### Data Collection

Consistent with the University of Illinois at Chicago Institutional Review Board, participants read and signed consent forms and were screened for health status to assure compliance with the exclusion criteria (i.e., brain injury, chronic bronchitis, smoking more than one cigarette per day). Following consent and screening procedures, three gel-electrodes were placed on the chest to record ECG and the participant was fitted with a LifeShirt® (Vivometrics™) to monitor respiration parameters and heart rate. The LifeShirt® is made of a stretchable fabric with two embedded inductance plethysmography bands at the thoracic and abdominal levels, which accurately measures continuous changes in tidal volume and provides accurate measures of beat-to-beat heart rate (Heilman & Porges, 2007). In the hospital setting, clinical staff set up an intravenous apparatus to deliver an infusion of saline or glycopyrrolate. Data are reported from the initial 5-minute baseline in the research laboratory and a 5-minute post-baseline monitored 45 minutes following a protocol involving a sequence of psychological tests evaluating affect recognition and auditory processing. The same experimental protocol (i.e., psychological tests) was administered in both research settings. In the hospital setting, the effect of vagal blockade was assessed during a 5-minute pre-infusion baseline and approximately 45 minutes following infusion during a 5-minute post-infusion seated baseline.

Vivometrics™ data acquisition software generated synchronous time series of calibrated tidal volume sampled 50 Hz and sequential heart periods timed to the nearest millisecond. The ECG waveform is sampled at 200 Hz by the LifeShirt® and then interpolated by Vivometrics™ software for identification of the R-wave peak with an accuracy equivalent to approximately  $\pm 2$  msec (Heilman & Porges, 2007). The heart period time series were visually inspected and missed R-wave detections and errors were corrected with CardioEdit (Brain-Body Center, Chicago, IL). The edited heart period sequences were analyzed with specific algorithms developed to quantify three metrics of RSA: Porges-Bohrer time domain method ( $RSA_{P-B}$ ), peak-to-trough (P2T), and the accumulated variances across a high frequency band (HF) from the spectral analyses.

### Data Analysis: RSA Quantification

**Porges-Bohrer Method ( $RSA_{P-B}$ )**—The Porges-Bohrer method assumes that heart period time series reflect the sum of several component time series. Each of these component time series may be mediated by different neural mechanisms and may have different statistical features. The Porges-Bohrer method applies an algorithm that selectively extracts RSA, even when the periodic process representing RSA is superimposed on a complex baseline that may include aperiodic and slow periodic processes. Since the method is designed to remove sources of variance in the heart period time series other than the variance within the frequency band of spontaneous breathing, the method is capable of accurately quantifying RSA when the signal to noise ratio is low.

The Porges-Bohrer method incorporates three strategic procedures. Each procedure is statistically justified and has a specific objective. First, the variance in the heart period time series associated with activity slower than spontaneous breathing, such as slow quasiperiodic processes often associated with vasomotor and blood pressure oscillations and nonstationary aperiodic trend, is removed with a moving polynomial filter (Porges, 1985). Second, to stabilize the estimates of RSA and to minimize brief aberrations in the amplitude of RSA that violate stationarity, RSA estimates are calculated sequentially over several short

time epochs (e.g., 30 seconds) to produce multiple estimates that can be averaged. Third, to ensure that the distributions of RSA estimates are normal, each epoch estimate is transformed by its natural logarithm (Riniolo & Porges, 2000).

RSA<sub>P-B</sub> analyses were conducted with CardioBatch, a program that implements the Porges-Bohrer method (Brain-Body Center, University of Illinois at Chicago). CardioBatch provides specific parameters for analyzing RSA in infants, adolescents, adults, exercising adults, and small mammals (e.g., prairie voles). The data presented in this paper were calculated with the standard adult parameters that define RSA across a band of frequencies (i.e., .12–.40 Hz) associated with spontaneous breathing. The sequential procedures incorporated in CardioBatch are described below.

The time series of sequential heart periods was re-sampled at 2 Hz and processed with a 21-point cubic polynomial filter (zero mean, 3<sup>rd</sup> order) with a low-pass cutoff frequency of 0.095 Hz (Porges & Bohrer, 1990). The filter was convolved with the re-sampled time series to create a template of the low frequency and aperiodic trend components in the time series. The trend component was subtracted from the re-sampled heart period time series to create a stationary zero mean time series that is statistically appropriate for filtering with a finite impulse response filter. The resulting time series has a periodic oscillation at the frequency of spontaneous respiration (Denver, Reed & Porges, 2007). The removal of aperiodic trend and low frequency activity provides an opportunity to more accurately quantify RSA, which may be masked or distorted by slower activity and aperiodic trend (Porges & Bohrer, 1990). A Chebychev type I bandpass filter was applied to the residual series to remove any variance outside the bandwidth of spontaneous respiration (0.12–0.40 Hz). Given the roll-off of the filters used, variance associated with oscillations slightly outside the filter bandwidth will be partially passed through to the final filtered series. The final filtered series was divided into 30-second epochs, the variance of each epoch is transformed with a natural logarithm ( $\ln(\text{ms}^2)$ ), and the final estimate of RSA is the mean of the epoch values. The logarithmic transformation reduces skewness and kurtosis of RSA<sub>P-B</sub> to enable the data to more closely conform to the assumption of normality. By using the average of shorter log transformed epochs, the impact of any aberrant distortion is minimized. In contrast, transitory distortions may have a massive influence on other time and frequency domain methods for quantifying RSA in which component variations are additive across longer periods.

**Peak-to-trough algorithm (P2T)**—The peak to trough method measures the statistical range in ms of the heart period oscillation associated with synchronous respiration. Operationally, subtracting the shortest heart period during inspiration from the longest heart period during expiration produces an estimate of RSA during each breath. The peak-to-trough method makes no statistical assumption or correction (e.g., adaptive filtering) regarding other sources of variance in the heart period time series that may confound, distort, or interact with the metric such as slower periodicities and baseline trend. Although it has been proposed that the P2T method “acts as a time-domain filter dynamically centered at the exact ongoing respiratory frequency” (Grossman, 1992), the method does not transform the time series in any way, as a filtering process would. Instead the method uses knowledge of the ongoing respiratory cycle to associate segments of the heart period time series with either inhalation or exhalation.

Custom software designed in MATLAB (P2T) automated calculations and output mean peak-to-trough in ms, total number of breath cycles analyzed ( $M = 46.98$ ,  $SD = 16.44$ ), and the number of breath cycles from which RSA could not be calculated ( $M = 3.43$ ,  $SD = 6.73$ ). The algorithm is similar to the software described by Uijtdehaage at UCLA (1993) and is consistent with the procedures described by Grossman, van Beek, and Wientjes (1990) and the commercially available programs from James Long Company and Vivometrics™. By

providing an opportunity to inspect and to identify the synchronous heart rate and respiratory data, the P2T software provides an opportunity not available in the commercially available programs to make decisions regarding the inclusion in the analyses of respiratory cycles in which RSA could not be detected.

P2T re-sampled the heart period time series at 50Hz to match the sampling rate of the synchronous tidal volume data provided by the Vivometrics™ software. Each breath cycle was identified by the slope of the tidal volume curve. The mean respiration period between breath cycles was calculated in ms. Cross correlation was used to calculate the lag of greatest correlation between the respiration and heart period time series within a limit of one and a half breath cycles (i.e. 150% of the mean respiratory period for that subject). The heart period time series was shifted back by this lag to synchronize the respiratory and heart period oscillations. Respiratory cycles, in which either the inhalation or exhalation components were shorter than 20% of the mean respiratory period, were excluded from analysis. For all other cycles the peak-to-trough difference in heart period was recorded in ms and the mean of these range scores reported as the estimate of RSA. In some publications (Grossman, van Beek, & Wientjes, 1990; Grossman, Brinkman, & De Vries, 1992; Ritz, Thons, Fahrenkrug, & Dahme, 2005; Ritz, 2009) peak-to-trough estimates include all respiration cycles and assign a zero to the respiration cycles in which there are either no associated heart period change or an inverse relationship (e.g., heart period minima during exhalation instead of inhalation). In the current study the results of the analyses were similar whether or not the peak-to-trough estimates were calculated with or without the zero scores. Since assigning or not assigning a zero resulted in similar results and made the distribution of P2T more non-normal, we report analyses only for the P2T metric that excluded breath cycles (i.e. did not include zero scores).

Additional analyses evaluated two modified versions of P2T methodology. First, a natural logarithm transformed the P2T ( $\ln P2T$ ). Second, the P2T estimate was calculated and logarithmically transformed following detrending with the moving polynomial applied to quantify  $RSA_{P-B}$  ( $\ln P2T_{mpf}$ ). These quantitative manipulations were conducted to contrast the P2T methodology and  $RSA_{P-B}$  methods, when the data have been similarly detrended (i.e., mpf) and logarithmically transformed (i.e., ln). To ensure an appropriate comparison among the methods in these analyses, the P2T metric (a measure of range) was transformed into a measure of variance. This was accomplished by calculating an estimate of the amplitude of the periodic heart rate oscillation by dividing the P2T metric by two and then inserting the amplitude estimate into the formula for calculating variance from a periodic process (i.e.,  $A^2/2$ ).

**High frequency Heart Rate Variability (HF)**—Spectral analyses were conducted in MATLAB using functions provided in the signal processing toolbox. Heart period time series were re-sampled at 2 Hz, linear trend removed, and a 512-point Fast Fourier transformation (FFT) decomposed the time series into constituent frequency components. The spectral decomposition was not influenced by length of the time series and the spectrum was smoothed with a 15-point Parzen window. RSA was quantified by accumulating the spectral densities within the frequency band defined by the respiration frequencies (i.e., 0.12 – 0.40 Hz). This high frequency (HF) band was consistent with the 50% roll-off of the filters used in the  $RSA_{P-B}$  method described above. Additional analyses modified HF by applying a natural logarithm transformation to both the HF estimate ( $\ln HF$ ) and to the HF estimate calculated following detrending ( $\ln HF_{mpf}$ ) with the moving polynomial applied to quantify  $RSA_{P-B}$ . These transformations were conducted to contrast the spectral metric with  $RSA_{P-B}$ , when the data have been similarly detrended (i.e., mpf) and logarithmically transformed (i.e., ln).

## Results

### Correlations among RSA metrics

As illustrated in Table 1, the three metrics are highly inter-correlated during the baseline condition in both laboratory settings and all metrics are significantly correlated with heart period.

### Test-retest reliability of RSA metrics

The stability of the RSA metrics was evaluated by correlations between the two test sessions. The metrics were significantly correlated between sessions ( $RSA_{P-B} = .617$ ,  $P2T = .320$ ,  $HF = .532$ ). The test-retest correlation was significantly lower for P2T than for HF and  $RSA_{P-B}$ .

The internal stability of the RSA metrics was evaluated by partitioning the heart period time series collected in the office laboratory into two equal segments and correlating the RSA values between the two within session segments. The initial baseline from the less demanding office session, which did not require an infusion, was used to evaluate internal stability. All three metrics exhibited high within session correlations. The correlations between the two segments and between each segment and the total recording are presented in Table 2.

### Test-retest reliability with short epochs

The  $RSA_{P-B}$  method provides estimates of RSA during short sequential time periods. In the current study data were collected during a steady state baseline in which vagal regulation of the heart was assumed to be relatively constant. These experimental conditions provide an opportunity to evaluate whether very short epochs convey the same information as a longer baseline. To evaluate the relation between very short epoch and the entire baseline, correlations were calculated between each sequential short epoch (either in 1-minute or 30-second epochs) and the entire 5-minute baseline. As illustrated in Table 3, each short epoch is highly correlated with the entire baseline.

### Distributional characteristics

As reported in table 4, respiration rate was consistently within the .12–.40 Hz frequency band (i.e.,  $\pm 3$  sd) used to define RSA in the  $RSA_{P-B}$  and HF metrics. Only 0.5% of the total individual breaths within the entire database, were outside this defined frequency band. All breath cycles outside the bandwidth of the filter were sufficiently close to the cutoff frequencies to have virtually no influence on the calculation of the frequency band dependent RSA metrics (i.e., HF,  $RSA_{P-B}$ ) either because of the roll off of the digital filters or due to the asymmetry of respiration in which all respiratory cycles in the data set had either an inspiratory or an expiratory phase within the selected frequency band. In addition, note that the distribution of heart rate deviates from normality and provides a statistical justification for the use of heart period in the parametric analyses described in this paper.

### Assumptions for parametric analyses

The assumption that variables are normally distributed is a fundamental assumption for parametric statistical tests. Thus, the investigation of within or between subjects correlations, treatment effects or interactions in both univariate and multivariate analyses of variance, requires that the distributions conform to the assumption of normality (i.e., skewness and kurtosis less than  $\pm 2$ ). Although the metrics are highly correlated, the shapes of their distributions differ. In table 4, skewness and kurtosis are reported for each

variable. Of the three RSA metrics, only  $RSA_{P-B}$  had values of skewness and kurtosis within the bounds that define normality.

In addition to normality, if RSA metrics are used in repeated measures parametric analyses, including within subject regression analyses, they must conform to more stringent assumptions of homoscedasticity and sphericity. Violations of these assumptions lead to increased Type I and Type II errors (e.g., Huynh & Feldt, 1970). Parametric analyses assume homogeneity of variance (homoscedasticity), which would require that variations of RSA for subjects with high and low amplitude RSA be similar. In repeated measures designs constant covariance (sphericity) across a wide range of manipulations is assumed. However, features of experimental protocols used in psychophysiological research are likely to result in violations of the sphericity assumption. For example, the variances associated with autonomic responses contrasting a seated baseline to an exercise condition or a seated baseline to a vagal blockade condition are likely to violate the sphericity assumption, while the variances associated with contrasting a seated baseline to a period of psychological testing may not. In general, if the fundamental assumption of normality is not violated, then violations of both homoscedasticity and sphericity may be adjusted with a variety of techniques that reduce the degrees of freedom in the model to make significance tests more conservative (e.g., Greenhouse-Geisser correction). An alternative approach to adjusting the degrees of freedom in the ANOVA model is to use a multivariate analysis of variance (MANOVA), which assumes normality without requiring the assumptions of sphericity or homoscedasticity.

The assumption of homoscedasticity was tested with Box's M (Box, 1949; Anderson, 1958; Seber, 1984) during minimal subject demands contrasting seated pre- and post-baselines in the laboratory component of the study during which the participants were involved in a series of psychological tests. Although all subjects experienced the same procedures, data were partitioned into two groups based on the treatments they were to receive during the infusion condition (saline, glycopyrrolate) at the hospital setting. Even during minimal demands both P2T and HF exhibited large violations (see Table 5). These findings emphasize that during procedures common in psychophysiological research (e.g., sitting in a room without posture or exercise challenge) P2T and HF metrics are likely to violate the assumption of homoscedasticity. Thus, if statistical analyses dependent on the assumption of homoscedasticity have been used to describe the responses of P2T and HF, then the statistical inferences derived from these studies need to be critically evaluated. Moreover, since P2T and HF violate the assumption of normality, then the preferred methods to correct for violations of homoscedasticity, by either adjusting degrees of freedom in univariate analyses of variance models (e.g., Greenhouse-Geisser) or electing to use multivariate analyses of variance models, cannot "rescue" P2T or HF since they can only make the significance tests more conservative.

During the complete protocol, homoscedasticity was challenged by exposing groups of subjects to variations in context and treatment (i.e., testing with and without infusion, during the infusion approximately 50% received saline and approximately 50% received glycopyrrolate). The profound vagal blocking effect of the glycopyrrolate, by selectively reducing the variance among subjects in this condition relative to the other non-blockade conditions, resulted in a violation of homoscedasticity (and sphericity) for all metrics. Thus, during conditions in which there are great changes in vagal regulation of the heart, such as during blockade or exercise, it is likely that the assumptions of homoscedasticity and sphericity will be violated, regardless of the RSA metric, and either correction procedures (i.e., Greenhouse-Geisser) or multivariate methods that are not dependent on this assumption should be used. However, these methods for dealing with violations of homoscedasticity and



sphericity are dependent on the data being normally distributed and of the three RSA metrics only the RSA<sub>P,B</sub> metric is normally distributed.

### Correlations between RSA metrics and respiratory parameters

Since the regulation of RSA is neurophysiologically and neuroanatomically integrated with the neural circuits that regulate respiration, RSA should exhibit a degree of linear dependence on respiratory parameters and during steady state conditions (e.g., baseline) respiratory parameters should be correlated with RSA. Across RSA metrics (see Table 6), faster breathing was negatively correlated with RSA and greater tidal volume was positively correlated with RSA. The statistical significance of the correlations between the respiratory parameters and RSA differed among the RSA metrics. P2T had the strongest correlations with the two respiratory parameters. The correlation between P2T and respiration rate represented approximately 15% shared variance and the correlation between P2T and tidal volume represented approximately 10% shared variance.

To estimate population parameters of these correlations and to minimize the potential influence of outliers in the sample on the regression analyses, data were bootstrapped. Bootstrapping is sampling with replacement (Manly, 1997). In the study of initial baseline levels 65 pairs of RSA and respiratory pre-baseline values measured in the office environment were sampled with replacement 10,000 times to generate distributions of correlation coefficients between the two respiratory parameters (i.e., respiration rate, tidal volume) and each RSA metric. From the distribution of the 10,000 correlations the median correlation was reported as a population estimate. The median correlations of the RSA metrics with the respiratory parameters are listed in Table 6. The sample correlations are virtually identical to the bootstrapped estimates reported.

Given the median bootstrapped correlations, the probability (as a function of sample size) of obtaining a significant correlation between respiration rate and each RSA metric are reported in Table 7. Based on the median bootstrapped correlations between P2T and respiration rate, only 27 participants are necessary to identify a significant effect 50% of the time, while much larger samples are necessary with the other methods. Similarly, smaller samples are needed with P2T than the other metrics to identify significant correlations between RSA and tidal volume. Since psychophysiological studies often report data collected from fewer than 40 participants, the findings relating RSA to respiration frequently have been reported with P2T, while the covariation has been more elusive with other metrics of RSA.

### Stationarity and the statistical features of RSA metrics

To investigate the influence of nonstationarity on the RSA metrics, the least stringent assumption for a “weakly” stationary time series (i.e., mean does not change over time) was evaluated. To conduct this analysis, the time sampled heart period time series was divided into equal length time bins. As a conservative measure of a violation of weak stationarity, a violation was identified as any significant difference in the means among the bins within the baseline time series (i.e. paired t-test). The baseline was incrementally divided into progressively smaller bins from approximately 150 seconds (i.e., 2 bins) to approximately 6 seconds (i.e., 50 bins). As illustrated in Figure 1, when the data were divided into 2 bins, with the P2T method 80% of the subjects had significant differences in mean heart period between the two bins. When the linear trend was removed from the baseline for the HF method, 17% of the subjects had significant differences in mean heart period between the two bins. When the moving polynomial was used to detrend, no subject had a significant difference in mean heart period between the two bins. As the number of bins increases to 5 (i.e., approximately 60 seconds), virtually all subjects using either the P2T or HF methods

violate the assumption. In contrast, no subject violates the assumption with the RSA<sub>P-B</sub> until the bins become smaller than 20 seconds (i.e., 15 bins within 300 seconds). Even when the bins shrink to less than 10 seconds and the number of contrasts increases, the moving polynomial effectively maintains a “stationary” mean for approximately 90% of the subjects.

Weakly stationary time series also assume that variance does not change over time. As a conservative measure of a violation of the constant variance assumption, based on Box’s rule of thumb for the more resilient between subject distributions (Box, 1953), a violation was identified as any ratio of variance exceeding 4.0 among the bins (described above) within the baseline. Violations were observed in every subject with bin sizes 10 seconds or longer for all data streams (i.e., time sampled heart period, linear detrended heart period, and moving polynomial detrended heart period). However, consistent with the procedures used in RSA<sub>P-B</sub>, when the variances were transformed with a natural logarithm there were no violations of this criterion. These analyses illustrate that the two assumptions of a weakly stationary time series are violated by the procedures used in the P2T and the HF metrics. In contrast, the methods embedded in RSA<sub>P-B</sub> provide statistical procedures that functionally “prestation” the data to enable estimates of RSA over short periods of time (see Bohrer & Porges, 1982). The above analyses illustrate the relative vulnerability of the specific methodologies to the assumption of weak stationarity.

### Sensitivity of RSA metrics to vagal influences

**Blockade**—A random subset of 25 of the 65 participants was administered an ACh blocking agent, glycopyrrolate, to investigate the sensitivity of each RSA metric and heart period to vagal blockade. Table 8 illustrates the profound effects on all of the cardiac metrics and the insensitivity of the respiratory parameters to the glycopyrrolate infusion. Note that heart period and RSA<sub>P-B</sub> had noticeably higher F-values than the other two RSA metrics.

To evaluate effect size, Glass’s  $\Delta$  was calculated for heart period and each RSA metric (Glass, 1976; Glass, McGaw & Smith 1981). To test for significant differences in effect size, confidence intervals for Glass’s  $\Delta$  were estimated with bootstrap analyses. This effect size measure is closely related to Cohen’s  $\delta$ , except that the pre-treatment variance is used as the estimate of the population variance. Glass’s  $\Delta$  was calculated on 10,000 samples of 25 subjects drawn with replacement from the original 25 individuals, who received the glycopyrrolate infusion. MATLAB code provided in the signal processing toolbox was used to estimate the 95% bias corrected and accelerated confidence interval of  $\Delta$  from the equation below. Due to profound effect of glycopyrrolate on the autonomic nervous system, the distribution of post-infusion measures was not representative of the population variance of each metric. Therefore, the standard deviation of the pre-infusion baseline was used in calculating Glass’s  $\Delta$ , as follows:

$$\Delta = \{ \mu_{\text{base}} - \mu_{\text{post}} \} / \sigma_{\text{base}}$$

As illustrated in Figure 2, the strongest effect size was observed in RSA<sub>P-B</sub>. In contrast, the effect sizes for P2T and HF were noticeable lower. In fact, the size of the effect in response to partial vagal blockade for RSA<sub>P-B</sub> was significantly greater than the other variables, including heart period. To further investigate the relative sensitivity of the RSA metrics, the partial eta-squared ( $\eta^2_p$ ) was calculated for each RSA metric in the complete experimental design. The complete design had three factors: 1) Group (saline or a glycopyrrolate infusion in the hospital session), 2) Setting (office building and hospital), and 3) Pre- Post baseline

(beginning and end of experimental session). Subjects were tested in both settings and randomly assigned to saline or glycopyrrolate infusion. Thus, the three-way interaction (Group  $\times$  Setting  $\times$  Pre/Post) provided an opportunity to observe how the massive changes in level and variability of the metrics due to vagal blockade, relative to the saline infusion and a non-infusion contrast, influenced the relative size of the effect as measured by  $\eta^2_p$ .

The three-way interactions are plotted in Figure 3 for each RSA metric and heart period. Note the unique response of  $RSA_{P,B}$  to glycopyrrolate in contrast to the stability of  $RSA_{P,B}$  during conditions when participants were not receiving the vagal blockade. The weaker effect sizes for P2T and HF are reflected in greater variability across settings and conditions. As illustrated in Figure 4, the  $\eta^2_p$  for  $RSA_{P,B}$  accounted for approximately three times as much variance as P2T and HF. The ratios among the three RSA metrics were consistent with the above analyses evaluating Glass's  $\Delta$  for the main effects of glycopyrrolate.

**Covariation of RSA metrics with changes in heart period**—Since the vagus provides the primary neural influence to the heart during non-exercise conditions, change in heart period during conditions that do not require a massive change in motor activity has been proposed as robust estimator of changes in cardiac vagal tone (Katona & Jih, 1975; Fouad, Tarazi, Ferrario, Fighaly, & Alicandri, 1984; Grossman & Kollai, 1993; Grossman, van Beek, & Wientjes, 1990). This argument is consistent with the above reported size of effects of heart period when evaluated with the pre-post glycopyrrolate infusion (i.e., Glass's  $\Delta$ ) and the three-way interaction from the complete design (i.e.,  $\eta^2_p$ ). Although heart period is sensitive to vagal influences, it also could be argued (see Porges, 2007), and supported by the above data on vagal blockade, that  $RSA_{P,B}$  is significantly more sensitive to vagal blockade than heart period. However, to enable a contrast among all RSA metrics and to be consistent with the above literature, change in heart period is used in the following analyses as a surrogate for changes in cardiac vagal tone. The saline infusion data were used for these analyses.

The contextual features of the saline infusion condition provided an opportunity to observe a large range of individual responses in RSA and heart period. Since the infusion condition did not necessitate an increase in activity or a change in posture that might result in transitory changes in sympathetic activation, the individual changes in heart period and RSA were assumed to be manifestations of changes in vagal influence to the heart. In contrast to the saline infusion, the glycopyrrolate infusion resulted in a massive decrease of the peripheral vagal efferent influences on the heart independent of any change in central regulation of vagal efferent activity that might be coupled to respiratory parameters.

In response to the saline infusion, changes in RSA were significantly correlated among the three metrics (see Table 9). The change in heart period was significantly correlated with changes in the HF and  $RSA_{P,B}$  metrics, but not P2T (see Table 10). These results suggest that changes in P2T are, of the three RSA metrics, the least sensitive to the changes in heart period assumed to be mediated by vagal mechanisms. The literature suggests that controlling respiratory parameters by pacing or via statistical adjustment will improve the covariation between P2T and vagal influence on the heart (Wilhelm, Grossman & Coyle, 2004).

### **Is the relation between the changes in RSA and heart period moderated by changes in respiration?**

Moderation analysis statistically defines under which conditions two variables are associated with one another as a function of a third variable (Kraemer, Wilson, Fairburn, & Agras, 2002). Thus, moderation analysis can be used to define the nature of a significant interaction effect of the influence of changes in respiratory parameters on the association between change in RSA and change in heart period. To evaluate the possible interaction between

changes in respiratory parameters and changes in RSA in predicting changes in heart period, multiple linear regressions were calculated using the centered scores for change in the RSA metric and changes in either respiration rate or tidal volume to predict changes in heart period. Centered scores reduce non-essential collinearity when computing interaction terms, while maintaining distributional parameters with a zero mean (Aiken & West, 1991). By assigning  $\Phi$ , as the change in a respiratory parameter (rate or volume), a general multiple linear regression formula for moderation may be stated as:

$$\Delta\text{Heart Period} = \beta_1 * (\Delta\text{RSA}) + \beta_2 * (\Phi) + \beta_3 * (\Delta\text{RSA} * \Phi) + C$$

In the first block of the regression analyses, the centered scores for change in RSA ( $\beta_1 * \Delta\text{RSA}$ ) and change in either respiration rate or tidal volume ( $\beta_2 * \Phi$ ) is entered. To test for the moderation, the interaction of change in RSA with a change in either respiration rate or tidal volume ( $\beta_3 * \{\Delta\text{RSA} * \Phi\}$ ) is entered in the second block. If the addition of the interaction term does not significantly improve the regression model, the main effect of change in RSA is not qualified by an interaction with  $\Phi$ . The main effect of change in RSA on change in heart period was qualified by a significant interaction with change in respiration rate only for P2T.

To further explore the moderation of changes in P2T by changes in respiration rate, the simple slopes of change in RSA on change in heart period for individuals who exhibited greater increases or decreases in respiration rate (distributions generated from values drawn at one standard deviation above and below the mean change in respiration rate) were tested following the method described by Holmbeck (2002). The slope differences of the simple regressions between the predictor variable (change in RSA) and the criterion variable (change in heart period) under different conditions of the moderator variable (i.e., greater decreases or increases in respiration rate) were contrasted. For these analyses, two distributions are generated from the centered (zero mean) distribution of absolute changes in respiration rate used in the above analysis. The new distributions for greater or lesser change in respiration rate maintain the higher order parameters (variance, skewness, and kurtosis) of the original change in respiration rate data. The other variables in the model, change in RSA (centered) and change in heart period, are not changed. For each distribution (greater change or less change in respiration rate), a new interaction term,  $(\Delta\text{RSA}) \times (\Delta\text{Respiration Rate})$  is computed. The multiple linear model is then computed by least squares regression to estimate the direct effect of change in RSA in predicting change in heart period for a distribution of have greater increases or decreases in respiration rate.

The hypothetical “Greater Change” and “Lesser Change” in respiration rate lines are plotted in Figure 5. The “Lesser Change” line represents no change in breathing rate from pre-to post-infusion. It is only when there is little or no change in breathing rate between pre and post is the change in P2T related to the change in heart period. However, as the difference in respiration rate between pre- to post-infusion increases, changes in P2T are not directly related to changes in heart period. The significance tests for the interaction term ( $\beta_3$ ), described in the general formula above are presented in Table 11. Changes in  $\text{RSA}_{\text{P-B}}$  were not qualified by changes in either respiration rate or tidal volume. Only for HF, did the interaction with tidal volume change approach statistical significance. The vulnerability of P2T to the moderation effect of changes in breathing rate confirms the necessity to pace breathing in order to ensure that changes in P2T are related to changes in cardiac vagal tone (Grossman, Karemaker & Wieling, 1991). However, there is no evidence that paced breathing would be required for  $\text{RSA}_{\text{P-B}}$ , since  $\text{RSA}_{\text{P-B}}$  is not moderated by variations in breathing rate.

**RSA metrics are not equivalent: Moderation by respiration and trend**—The above analyses provide strong documentation that  $RSA_{P-B}$  is more sensitive to vagal blockade and the relation between changes in  $RSA_{P-B}$  and changes in heart period during saline infusion is not moderated by respiration. Since  $RSA_{P-B}$  is consistently a more sensitive surrogate for cardiac vagal tone than the other RSA metrics, analyses were conducted to determine if the differences among the metrics could be, in part, explained by moderation by respiratory parameters and trend.

To address this question, moderation analyses evaluated possible interactions between respiratory parameters and either P2T or HF in predicting  $RSA_{P-B}$  during baseline in the office laboratory. Although the P2T and HF were significantly correlated with  $RSA_{P-B}$ , the main effect was qualified by significant interactions with both respiratory rate and tidal volume for P2T and HF. The significance tests for the interaction term ( $\beta_3$ ), described in the general formula above are presented in Table 12. The respiratory parameters had no significant main effects in the regression models. Similar interaction effects by respiratory parameters were observed when logarithmic transformations of the P2T and HF metrics were used to predict  $RSA_{P-B}$ .

For respiration rate, the sample means for these distributions were 1 standard deviation ( $\pm 0.054$  Hz, see Table 4) above and 1 standard deviation below the mean respiration rate (i.e., 0.22 and 0.32 Hz). The regression weight for RSA in this model is used to calculate the simple slopes.  $RSA_{P-B}$ , as a function of the P2T and HF, are plotted at  $\pm 1$  standard deviation. For example, with the uncorrected P2T this is equivalent to 31.10 and 167.92 ms (see Table 4).

To visualize the significant interaction of respiration rate with the P2T and HF, the simple slopes of these RSA metrics on  $RSA_{P-B}$  are plotted in Figures 6 and 7. In both figures, the slopes of the two regression lines are statistically different. In the fast breathing distribution, the steep slope indicates that the P2T and HF metrics significantly predict  $RSA_{P-B}$ . The shallow slope for the slow breathing distribution indicates that the P2T and HF metrics are poor predictors of  $RSA_{P-B}$  for slowly breathing individuals.

We further investigated the influence of trend and whether the degree of violation of stationarity moderated the relation between P2T or HF and  $RSA_{P-B}$ . As an exploratory analysis the baseline heart period time series was divided into 15 equal segments and paired t-tests were used to identify significant differences among pairwise contrasts of all the segments. From the set of significantly different means, we took the average absolute difference in ms as an estimate of the magnitude of deviation from a stationary baseline. A greater mean difference was assumed to reflect a greater violation of stationarity.

Moderation analyses were performed to see if nonstationarity of mean moderated the relation between P2T or HF and  $RSA_{P-B}$ . As illustrated in table 13, the analyses demonstrated a significant interaction between this component of trend and P2T and HF magnitude in the prediction of  $RSA_{P-B}$ . Only when the data are relatively stationary is there linear relationship between either P2T or HF and  $RSA_{P-B}$ . Similar to the above respiration moderation analyses, logarithmic transformations did not remove the moderation effects of trend. Thus, emphasizing the moderation effects of both respiration and trend are not solely determined by distributional features (i.e., conforming or violating assumptions for parametric analyses) and still exist when the data are logarithmically transformed.

Interestingly, the linear correlations between the nonstationarity measure and respiration rate and tidal volume were not significant ( $r = -0.14$ ,  $p = 0.283$  and  $r = -0.05$ ,  $p = 0.676$ , respectively), indicating that the nonstationarity of the heart period time series was not

related to the respiratory parameters. An interpretation of these results is that trend interacts with the RSA measures P2T and HF in the prediction of RSA<sub>P-B</sub> and that the magnitude of this error is a function of the subject's respiration pattern.

### The influence of detrending and logarithmic transforms on P2T and HF

Based on the above analyses, it appears that two quantitative steps embedded in the RSA<sub>P-B</sub> method may be responsible for the enhanced distributional features and the removal of moderation by respiration and nonstationarity. These features are: 1) dynamic detrending with a moving polynomial to remove sources of variance slower than the signal of interest, and 2) logarithmic transformation of the variance estimate.

Logarithmic transformations are useful in transforming metrics of variance to improve distributional characteristics. Periodic processes can be conceptualized as variance measures by the following formula in which  $A$  represents the amplitude of the signal:  $A^2/2 = VAR$ . Since spectral densities are a measure of variance, these values can be easily transformed. However, the P2T method produces in a measure of range or twice the amplitude of the signal. Thus, prior to logarithmic transformation, the P2T values were divided by 2 to produce measures of amplitude and then inserted in to the above formula for variance. To further evaluate comparisons among the methods, an additional value for RSA<sub>P-B</sub> was calculated over the entire condition and not averaged over repeated 30 second epochs. As illustrated in Table 14, following logarithmic transformation the three RSA metrics have values of skewness and kurtosis within the bounds that define normality.

When the moving polynomial detrending algorithm was used to preprocess the data prior to applying the P2T or HF algorithms and the variances estimates from these metrics were logarithmically transformed, the statistical features and sensitivity of all metrics converged. Removing trend with the moving polynomial prior to calculating both HF and P2T and then applying a natural logarithmic transformation to the extracted variances (i.e.,  $\ln HF_{mpf}$ ,  $\ln P2T_{mpf}$ ) successfully removed respiration rate and trend as moderators. Figure 10, illustrates that the size of effect in response to glycopyrrolate infusion is similar for all three RSA metrics, once they are preprocessed with the moving polynomial and post-processed with the logarithmic transformation. In the analyses illustrated in Figure 10, only one value for RSA<sub>P-B</sub> was calculated over the entire baseline rather than the preferred mean of short bins. This was done to limit the differences among the RSA metrics to only the method used to calculate variance (i.e., min-max for the P2T, spectral band for HF, and bandpassed variance for RSA<sub>P-B</sub>). These analyses illustrate that the vulnerabilities described in the above sections of the various RSA metrics are not inherent in spectral or peak-to-trough methods, but are due to an inappropriate application of the methods. The analyses emphasize that the observed differences in RSA metrics are a function of the methodology applied to isolate RSA from other sources of variance in the heart period time series and whether the metric conforms to assumptions necessary for parametric analyses.

## Discussion

### Are commonly used RSA metrics equivalent?

Porges (2007) proposed that several assumptions regarding RSA are based on claims that have not been appropriately tested with sufficiently rigorous methods. Two of these assumptions are challenged in this paper: 1) highly intercorrelated RSA metrics are equivalent, and 2) RSA metrics need to be statistically adjusted for ventilatory parameters to accurately estimate cardiac vagal tone.

In earlier research Grossman et al. (1990) reported that the three frequently used RSA metrics described in this paper were highly correlated. Our data are in agreement with this

finding. Because the various RSA metrics have high intercorrelations, it has been assumed that the metrics “could be considered almost interchangeable” and “RSA comparisons can be freely made from study to study, even when investigators employ different of these quantification procedures” (Grossman et al., 1990, p.713). Consistent with this assumption of equivalence, if one RSA metric is improved by statistically correcting for breathing rate, it would be assumed that all methods will improve. Inferences, based on an assumption of equivalence, regarding the dependence of RSA on respiration, have promoted an appearance of controversy when various corrections to RSA involving respiratory parameters are challenged. This paper poses the empirical question of whether RSA metrics are truly equivalent. If the metrics are not equivalent and behave differently in terms of statistical features and interactions with respiratory parameters and nonstationarities, then there is no controversy. In addition, if the removal of trend and a logarithmic transformation of the variance estimates of RSA results in all RSA metrics behaving similarly, then instead of controversy there will be a better understanding of the sources of error variance that contribute to the variations among specific RSA metrics.

### Statistical Features of RSA metrics

All metrics exhibited very high split-half correlations during baseline suggesting excellent short-term stability. When the interval between test sessions is expanded to a two-week interval, the magnitude of the test-retest correlations drops with P2T having significantly lower test-retest correlations than either RSA<sub>P,B</sub> or HF.

Table 15 describes the statistical features of the three RSA metrics and designates if, during baseline conditions, the RSA metric failed to conform to assumptions necessary for parametric analyses. The distributional features of P2T and HF do not conform to assumptions related to normality. In addition P2T and HF failed Box’s M test for homoscedasticity, when a pre-baseline versus a post-baseline was contrasted for two groups receiving the identical mild psychological tests. It is important to note that heteroscedasticity and sphericity violations can be accommodated by using Multivariate Analyses of Variance or corrections of the degrees of freedom in Analyses of Variance models (e.g., Greenhouse-Geisser correction). However, these methods require that the underlying distributions are normal and this is not the case for P2T and HF.

Distributions of variance measures, similar to the RSA metrics, often can easily be adjusted to approximate normality by the use of logarithmic transformations. Although the logarithmic transformation is a defining feature of RSA<sub>P,B</sub> and has frequently been applied to HF, it is rarely applied to P2T. The data from the Grossman et al (1990) study, which provides the only other comparison in the literature among the three metrics investigated in this study, has served as basis for the equivalence assumption. It is important to note the reported high intercorrelations among the metrics were based on logarithmically transformed variables. This important adjustment to the distribution of P2T values has not been generalized to current applications of the P2T methodology nor has it been mandated in defining HF.

### RSA metrics and respiration

Our findings confirm previous reports regarding the simple direct effect of respiratory rate and tidal volume on RSA magnitude. Grossman et al. (1990) evaluated the correlation between respiration and each of the three RSA metrics tested in the current study. They reported that P2T had the highest correlations with respiration rate. Since Grossman et al. (1990) focused on within-subject correlations, a direct comparison with our data is not possible. Nonetheless, the pattern is consistent with our findings that the P2T metric had the

highest correlations with respiration rate among the three metrics and was the only RSA metric correlated with tidal volume.

### Vulnerability of RSA metrics to nonstationarity

Previous publications have addressed the vulnerability of the P2T metric to distortions by trends, slow periodic cycles, lack of symmetry in periodic oscillations, and signal to noise ratio (e.g., Byrne & Porges, 1993; Porges & Byrne, 1992; Weber et al., 1992). Most of these features can be discussed within the context of the violation of weak stationarity. Weak stationarity requires that the mean and variance of the time series do not vary across time. In the above referenced papers, analyses of real and simulated data carefully articulated the vulnerability of the P2T method to the naturally occurring changing features of trend, respiration rate, and the relative ratio of the amplitude of RSA to trend. Byrne and Porges (1993) provided documentation of the interactive influences of trend, respiratory rate, and RSA amplitude on the estimate of RSA derived from the P2T metric. Moreover, they detailed the important role of the signal to noise ratio in this vulnerability and acknowledged that, if RSA were of large amplitude relative to trend, the distortion of RSA with P2T would be negligible. From the Byrne and Porges (1993) perspective, the fundamental problem with the P2T method is that measurement error of the P2T technique cannot be specified a priori and held constant across any time series being processed. Instead, the accuracy of the method is a function of the dynamically changing trend and signal characteristics being used to define RSA (amplitude and frequency) within the data set being analyzed. Given these sources of variance statistical adjustments may be impossible without constraining the subject's respiratory parameters. Thus, it is logical that pacing respiration has been strongly recommended by advocates of the P2T methodology to deal with the influence of variations in respiration on this metric.

In contrast to documented statistical arguments of the problems inherent in the P2T metric, Grossman (1992) has argued that the P2T is “robust and relatively unaffected by various types of nonstationarity” and has continued to encourage the application of this methodology because it is easier to calculate and provides an estimate of RSA on a breath-by-breath basis regardless of the breathing rate (Grossman, Wilhelm, Spoerle, 2004). The data in this paper contradict these assumptions.

The data analyses described in this study provide an additional opportunity to document, with real data and not simulations, the vulnerability of the P2T and the other RSA metrics to violations in stationarity. In this paper the question was approached by investigating separately the two assumptions, constant mean and constant variance, required for a weakly stationary time series. First, analyses documented the moderation effect of nonstationary mean levels on the estimates of RSA derived from both the P2T and HF methods. The findings illustrate that the P2T metric is significantly influenced by nonstationarity. Moreover, linear detrending, which is assumed to minimize nonstationary shifts in mean, is relatively ineffective in isolating the HF metric from trend. In contrast, due to the effectiveness of the moving polynomial in removing nonstationary trend,  $RSA_{P-B}$  was not influenced by nonstationary mean levels.

Constant variance, the second assumption for weak stationarity, was also investigated. Most psychophysiological research is based on the sensitivity of variables to contextual or neurophysiological features and manipulations. In this case, the variable of interest, RSA, is a measure of variance. When variables are defined by variance, the assumption of a constant variance needs to be interpreted cautiously, especially since experimental manipulations are intended to systematically change the magnitude of the variance estimate. In our analyses we have attempted to evaluate the assumption of constant variance within the constraints of how each variable is defined. Only the  $RSA_{P-B}$  metric maintained a ratio of variances



among the bins (see results) within the criterion for constant variance. The  $RSA_{P-B}$  metric complies with this assumption, since the methodology incorporates a logarithmic transformation of the variance. In addition, the logarithmic transformation enables the  $RSA_{P-B}$  metric to conform to assumptions regarding constant variance and distributional features required for the parametric analyses.

The modification of the P2T metric by inserting zeros in the calculation when a heart period min-max cannot be associated with a respiratory cycle has been recommended (Grossman & Svebak, 1987), although the effect of this recommended manipulation on statistical features has not been tested. The data in this study provided an opportunity to make this evaluation. With the current data, the effect of inserting zeros had negligible effects on the metric. For example, inserting zeros did not influence any of the features described in the summary table. When the two measures of P2T (with and without zeros) were correlated, the correlation approached unity ( $r = 0.997$ ). Overall the difference between the two values was small ( $M = 2.87$  ms,  $SD = 6.09$  ms) with the addition of zeros slightly lowering the estimate of RSA. Moreover, when the signal to noise ratio was enhanced (i.e., higher amplitude RSA relative to trend) by reducing the background trend with the moving polynomial filter, the number of zeros was reduced by 50%. In addition, as further documentation that the P2T metric is vulnerable to nonstationary levels, there was a significant correlation between the degree of nonstationarity (i.e., greater differences between the means of the bins as described in the moderation analyses above) and the number of zeros assigned ( $\rho = 0.443$ ,  $p < 0.001$ ). When the data were more nonstationary, more zeros were assigned. The dependence of identifying zeros on trend has been previously described by Byrne and Porges (1993), who illustrated with simulations that a “real” peak can be lost in trend and that peak can be time shifted outside the defined rules for identifying a peak or a trough. Thus, although the assigning of zeros has been proposed as a method to adjust the P2T metric for the influence of trend and baseline nonstationarities, the data from this study demonstrate that assigning zeros has no effect on the ability to extract the RSA signal from trend.

### Sensitivity to vagal mechanisms

Specificity of measurement may translate into superior sensitivity when the three RSA metrics are contrasted on pre-post blockade data. As the relative error variance in RSA measures increases, the observed effect sizes in response to a direct manipulation on cardiac vagal tone (i.e., glycopyrrolate infusion) are reduced. The relative sensitivity of each RSA metric can be compared by calculating effect size, Glass’s  $\Delta$ , and confidence intervals for each metric in response to blockade. As illustrated in Figure 2,  $RSA_{P-B}$  was significantly more sensitive than the other RSA metrics to a partial blockade of efferent vagal nerve traffic. In addition, consistent with an earlier study (Porges, 1986),  $RSA_{P-B}$  was more sensitive to cholinergic blockade than heart period. However, as illustrated in Figure 10, when P2T and HF were preprocessed with the moving polynomial filter and logarithmically transformed, all metrics were equally sensitive to vagal blockade. Similarly, when testing the complete experimental design across contexts with different infusion groups (blockade and saline), the RSA metrics performed differently. Consistent with blockade analysis, when the effect size for each RSA metric was estimated with  $\eta^2_p$ , the  $\eta^2_p$  for  $RSA_{P-B}$  was more than **three times greater** than either P2T or HF. These findings provide strong empirical support that  $RSA_{P-B}$  is more sensitive to vagal mechanisms than either P2T or HF. Moreover, these analyses provide a strong justification for proposing that  $RSA_{P-B}$  is a sensitive surrogate variable for cardiac vagal tone and substantially more sensitive to vagal influences than P2T or HF.

### **Is it necessary to pace breathing or statistically correct RSA metrics?—**

Because both respiration and RSA involve common brainstem circuits, no measure of RSA is free from covariation with respiratory activity. This does not preclude the utility of RSA amplitude as a surrogate measure of cardiac vagal tone. However, arguments have been made that if there is a correlation between RSA and respiration, it needs to be removed to use RSA as an index of cardiac vagal tone. Or, is it possible that if respiration parameters differentially influence the RSA metrics, then controlling for respiration may be effective in improving the sensitivity to vagal influences of one metric and not for the other metrics?

To evaluate this question we looked at the change in heart period in response to the saline infusion. The response to the saline infusion was selected because the saline infusion produced a relatively large range of individual differences in heart period reactions (i.e., from a decrease of 70 ms to an increase of 93 ms) without requiring posture or motor demands. Although our bias would be to propose that change in  $RSA_{P,B}$  is a better surrogate for cardiac vagal tone than change in heart period, we wanted to evaluate the moderation by respiratory parameters of the performance of each of the three RSA metrics in predicting a surrogate variable of cardiac vagal tone. Given this restriction, we evaluated the relationship between the response to saline infusion of each measure of RSA and heart period. The selection of changing levels of heart period as an index of vagal influences to the heart is consistent with the literature. For example, it has been proposed that, even during physical exertion, “minute-to-minute changes in heart rate over the day, largely due to fluctuations in physical exertion, may provide an independent index of cardiac vagal tone” (Grossman, Wihelm & Spoerle, 2004).

Moderation analyses demonstrated that the direct effect of change in P2T in predicting change in heart period was linear for subjects who maintained a constant respiration rate from a pre-baseline to a post-saline infusion baseline. For subjects who either increased or decreased respiratory rate across the experimental session, change in P2T was not directly related to change in heart period. Thus, the distortion of P2T by respiratory parameters cannot be remedied by correcting for the influence of respiration on P2T, as advocated by proponents of statistical corrections to RSA measures. Since the relationship between changes in P2T and changes in heart period varies as a function of the subject’s respiration rate and cannot be corrected by simple linear regression, the only option to “rescue” this variable is to pace the subject’s breathing. A dependence of P2T on constant respiration is consistent with the pacing or fixed respiration strategy proposed by Grossman et al. (Grossman, Karemaker & Wieling, 1991). Unfortunately, pacing not only limits the application of the P2T metric to laboratory situations, but also limits the tasks and conditions during which P2T can be monitored. Logarithmic transformations of P2T did not remove the moderation effect of respiration rate. The relation between changes in heart period and changes in either  $RSA_{P,B}$  or HF were not moderated by respiration rate.

These findings, seem to explain, in part, why advocates of the P2T methodological have argued that RSA is a poor measure of cardiac vagal tone that can be improved by keeping respiration rate constant (i.e., paced breathing) and by statistically adjusting for individual differences in tidal volume. They are correct, if their findings are based on P2T. These findings provide additional evidence of a lack of equivalence among RSA metrics. Thus, based on the data presented, the assumed dependence of RSA, as index of cardiac vagal tone, on respiratory parameters is a function of the method used to quantify RSA.

### **The equivalence of RSA metrics is moderated by respiration rate and trend—**

As an attempt to further explicate the inconsistency in the RSA literature on the role respiration plays in defining or adjusting RSA as a surrogate measure of cardiac vagal tone, the equivalence among the metrics was investigated with moderation analyses. As a strategy

to evaluate equivalence, moderation analyses were used to question equivalence among the metrics by evaluating whether the prediction of  $RSA_{P-B}$  by either P2T or HF was moderated by respiratory parameters. As illustrated in Figures 8 and 9, the moderation analyses demonstrated a linear prediction of  $RSA_{P-B}$  by P2T and HF only for fast breathers. With slow breathers the relation was orthogonal. Among slow breathers, individual differences in RSA magnitude for the P2T and HF metrics were not mapped into individual differences in  $RSA_{P-B}$ . The moderation analyses contribute to our understanding of why proponents of the P2T have argued for paced breathing, especially at slower breathing rates. The commonly used slow paced breathing would decouple the P2T and HF RSA metrics from the interaction with respiration rate (by fixing respiration rate at a single value), but pacing would not improve the inherently poor statistical features of these metrics.

In addition, the degree that the heart period time series has a nonstationary mean (i.e., trend) moderates the relationship between the P2T and HF metrics and  $RSA_{P-B}$ . The moderation analyses emphasize that when the data have a stationary mean, the relationship between either P2T or HF and  $RSA_{P-B}$  is direct, but when there is a nonstationary mean the relationship is distorted. Logarithmic transformation of both P2T and HF did not remove the moderation effects of either respiration rate or trend in predicting  $RSA_{P-B}$ . Thus, due to the interaction effects of both respiration rate and trend, the three metrics are not equivalent even when they are logarithmically transformed.

### Statistical modifications optimize P2T and HF

The  $RSA_{P-B}$  method applies two statistical features that set it apart from the other methods. First, it uses a logarithmic transformation. Second, it applies a moving polynomial filter to remove trend and periodic processes slower than spontaneous respiration. Although not mandated or recommended for either the P2T or HF metrics, the logarithmic transformation reduces problems associated with kurtosis, skewness, homoscedasticity, and corrections for violations of sphericity. When the logarithmic transformation is applied to both P2T and HF the issues regarding assumptions for parametric analyses are corrected, although it does not remove the interaction between these metrics and either respiration or trend.

Contributing to the inconsistencies in the literature, logarithmic transformations are frequently reported for HF and rarely reported for P2T. P2T is usually reported as min-max differences in milliseconds. When modifications (e.g., correcting for tidal volume) to the P2T are suggested, they are often administered to the min-max difference (e.g., Grossman & Taylor, 2007) and the statistical behavior of these new modified metrics are not known. In addition, in some publications, normalized units have been used to describe HF. The use of normalized units provides an index of relative variance of a frequency band (i.e., operationally defining RSA) relative to the total variance of the time series. Normalized units have been used to quantify ratios between the spectral densities accumulated in specific frequency bands. These ratios have been proposed as indicators of sympathovagal balance (e.g., Montano et al., 1994). This quantitative strategy precludes the use of RSA as an accurate measure of amplitude that could be compared between subjects or across testing sessions in within subject designs.

Moderation by respiration is not remediated unless the underlying sources of variance outside the frequency band of respiration are removed. Spectral analysis is compromised in performing this task when the data are nonstationary and include slow trends and aperiodic processes. Linear detrending or any low order polynomial fit over the entire time series will not map well into the slower oscillations and aperiodic sources of variance, which will fold into the spectral measures of RSA. The  $RSA_{P-B}$  method solves this problem by applying a moving polynomial filter that dynamically removes slow oscillations and aperiodic trends that are embedded in the beat-to-beat heart rate time series. This detrending technique is

critical in removing the moderator effects of both respiration rate and nonstationary trend. For example, both P2T and HF will approximate the behavior of the  $RSA_{P,B}$  metric, if modified to incorporate both a logarithmic transformation and preprocessed with the moving polynomial filter. When these modifications are applied to P2T and HF, similar to the  $RSA_{P,B}$  metric the modified variables will conform to parametric assumptions and will not interact with respiration rate or trend.

### Concluding comments

An understanding of neurophysiological mechanisms has informed the development of the signal processing steps involved in defining  $RSA_{P,B}$ . In developing  $RSA_{P,B}$  it has been assumed, based on the accepted identification of cardioinhibitory vagal fibers by their respiratory rhythm, that the functional impact of these inhibitory fibers would be observed as a periodic pattern in the beat-to-beat heart period time series. This periodic pattern would have two defining characteristics: 1) a periodicity associated with spontaneous breathing, and 2) an amplitude that would be related to the vagal efferent activity influencing the sinoatrial node. Thus, from a signal processing perspective the agenda would be to isolate and quantify this signal. It is clear that when the signal is isolated from trend and the metric is transformed to be consistent with parametric analyses, the metric describing this signal is extremely sensitive to manipulations of vagal influences and free from the moderating influences of trend and respiration rate.

The data from this study demonstrate that the RSA metrics differ on the statistical dimensions investigated in this paper. With this new understanding of the statistical features of specific RSA metrics, researchers may be better informed of the vulnerabilities of specific metrics and may agree on a standard metric to investigate both the validity of RSA as an index of cardiac vagal tone and the sensitivity of RSA to behavioral, psychological, and physiological manipulations.

The link between P2T and respiration rate has been thoroughly described in the literature. According to Byrne and Porges (1993), P2T will always perform better when RSA (i.e., the periodic respiratory “waveform” in the heart period time series) is reflected by a slope that is “steeper” (i.e., faster rise time) than the trend upon which the signal is superimposed. Information regarding the distortion of P2T by trend and respiration rate has not been incorporated in subsequent applications of P2T, nor has this information been used to improve the signal extraction features of P2T. Contrary to an acknowledgment of these weaknesses and the development of modifications of the P2T method or statements of limitations, this information has been ignored and has not influenced the application of the P2T methodology in psychophysiological research. For example, inconsistent with the documented vulnerabilities of P2T, it has been stated that “simulation studies indicate that this estimation procedure [P2T plus zeros when no peak or trough is identified] is robust against nonstationary trends within the human physiological range of heart rate” (see p. 729 Grossman, Wilhelm, and Spoerle, 2004).

Although all three metrics are highly inter-correlated, only the  $RSA_{P,B}$  conforms to the statistical assumptions necessary for parametric analyses. Unique to the three RSA metrics,  $RSA_{P,B}$  is not moderated by either respiration or the degree that the data violate the assumption of weak stationarity, nor is the correlation between change in  $RSA_{P,B}$  and change in heart period moderated by ventilatory parameters. Moreover, the sensitivity of  $RSA_{P,B}$  to partial vagal blockade is significantly stronger than the other two RSA metrics. Decomposing the methodology used to quantify  $RSA_{P,B}$  identified two features (i.e., detrending, logarithmic transformation) that when applied to both P2T and HF optimize statistical features, remove moderation by respiration, distortion by trend, and improve

sensitivity to cholinergic blockade. When these modifications are made to the P2T and HF, all metrics behave similarly.

The data illustrate that the basis for aggressive arguments to either pace breathing or to statistically correct breathing parameters from RSA may be due to the use of RSA metrics such as P2T, which both violate assumptions for parametric statistical analyses and are moderated by respiration and trend. In addition, these findings provide a plausible explanation of the reported dissociation between “cardiac vagal tone” and RSA when RSA is quantified with P2T (e.g., Grossman & Taylor, 2007). Regardless of these vulnerabilities, P2T continues to be used and has been implemented into turnkey programs (e.g., James Long Company).

The analyses demonstrate that, given current methodologies available to quantify RSA,  $RSA_{P-B}$  is the most appropriate measure of RSA. In addition, it could be argued, based on the above analyses and demonstrating the greater sensitivity to vagal blockade and stronger relations with changes in heart period, that  $RSA_{P-B}$  is the best surrogate for cardiac vagal tone.  $RSA_{P-B}$  is derived by a pragmatic approach that includes the following steps: 1) eliminating variance from the heart period time series outside the spontaneous breathing frequencies, 2) summing the variance that remains, and 3) normalizing the distribution of variances by the natural logarithmic transformation. These procedures result in a measure that is not moderated by respiration or trend, conforms to the assumptions required for parametric analyses, and is more sensitive to vagal blockade than heart period (i.e., the preferred criterion variable used in physiology). When the steps of detrending and logarithmic transformation are applied to P2T and HF, the modified metrics share the advantages of  $RSA_{P-B}$ . Other methods, such as CMetX (Allen, 2007), that incorporate procedures similar to those implemented in  $RSA_{P-B}$  metric (i.e., stepwise detrending over short periods of time to remove trend, extraction of the variance in the frequency band of spontaneous breathing, and logarithmic transformation of variance estimates) should perform with statistical features and sensitivity to vagal blockade similar to  $RSA_{P-B}$ .

Methodologies that quantify RSA with P2T and HF are functionally agnostic to complex trend and the influence of slower components. Both methods work extremely well when there is a large periodic signal and the background trend is of low amplitude. However, it has been demonstrated that trend and slower components can shift levels of the peak and trough or shift the timing of the peak and trough relative to breathing (see Byrne and Porges, 1993) with the P2T method. Moreover, HF and all applications of spectral analyses assume stationarity and violation of stationarity with complex trend and slow components may result in sources of variance outside the frequency band associated with respiration being attributed to HF. Linear detrending does not solve this problem.

The quantitative steps embedded in the  $RSA_{P-B}$  are an example of how knowledge of neurophysiology can inform quantitative procedures to extract a signal from the heart rate pattern that behaves consistently and robustly as a dynamic index of regulation of the heart via vagal pathways originating in the nucleus ambiguus. Neurophysiology informs us that the functional output of the myelinated vagus originating from the nucleus ambiguus has a respiratory rhythm. Thus, there would be a temporal relation between the respiratory rhythm being expressed in the firing of these efferent pathways and the functional effect on the heart rate rhythm manifested as RSA. From a signal processing perspective, there are two tasks: 1) defining the frequency band to extract the periodic signal, and 2) removing all sources of variance that might influence the ability to accurately describe the extracted variance in the designated frequency band. In the case of the heart period time series, the variance not associated with the periodic signal defined by breathing frequencies would be manifested as complex trend and slower periodic or quasi-periodic components. The  $RSA_{P-B}$  method was

designed to perform these two tasks. When these tasks are not effectively implemented, even if the distributions still conform to parametric assumptions (i.e., via logarithmic transformations), the RSA metric (e.g., P2T) is a poor index of cardiac vagal tone and the relation with indices of cardiac vagal tone is moderated by trend and respiratory parameters.

In summary, the data reported demonstrate that although the three commonly metrics of RSA are highly inter-correlated, the metrics differ in terms of statistical features, moderation by respiration, distortion due to nonstationarities, and sensitivity to vagal manipulations. Thus, the analyses confirm that the metrics are not equivalent and that  $RSA_{P-B}$  is the most sensitive metric to vagal influences. The net result of these findings is that the literature reported during the past 30 years needs to be revisited if either P2T or HF metrics were used without appropriate transformations and effective detrending procedures.

### Highlights

- Commonly used metrics for quantifying RSA are highly correlated.
- Several RSA metrics violate distributional requirements for parametric analysis.
- The Porges-Bohrer metric ( $RSA_{P-B}$ ) is appropriate for parametric analyses.
- $RSA_{P-B}$  is not moderated by respiration or nonstationarity in heart period.
- $RSA_{P-B}$  is significantly more sensitive to vagal blockade than other common metrics.

## Acknowledgments

The project described was supported, in part, by Award Number R01HD053570 from the Eunice Kennedy Shriver National Institute of Child Health & Human Development and by training grants T32 MH067631 and T32 MH18882 from the National Institute of Mental Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

### Special Thanks

The authors would like to thank Prof. Linda J. Skitka for her valuable instruction and guidance in developing the multiple linear regression models used in this report.

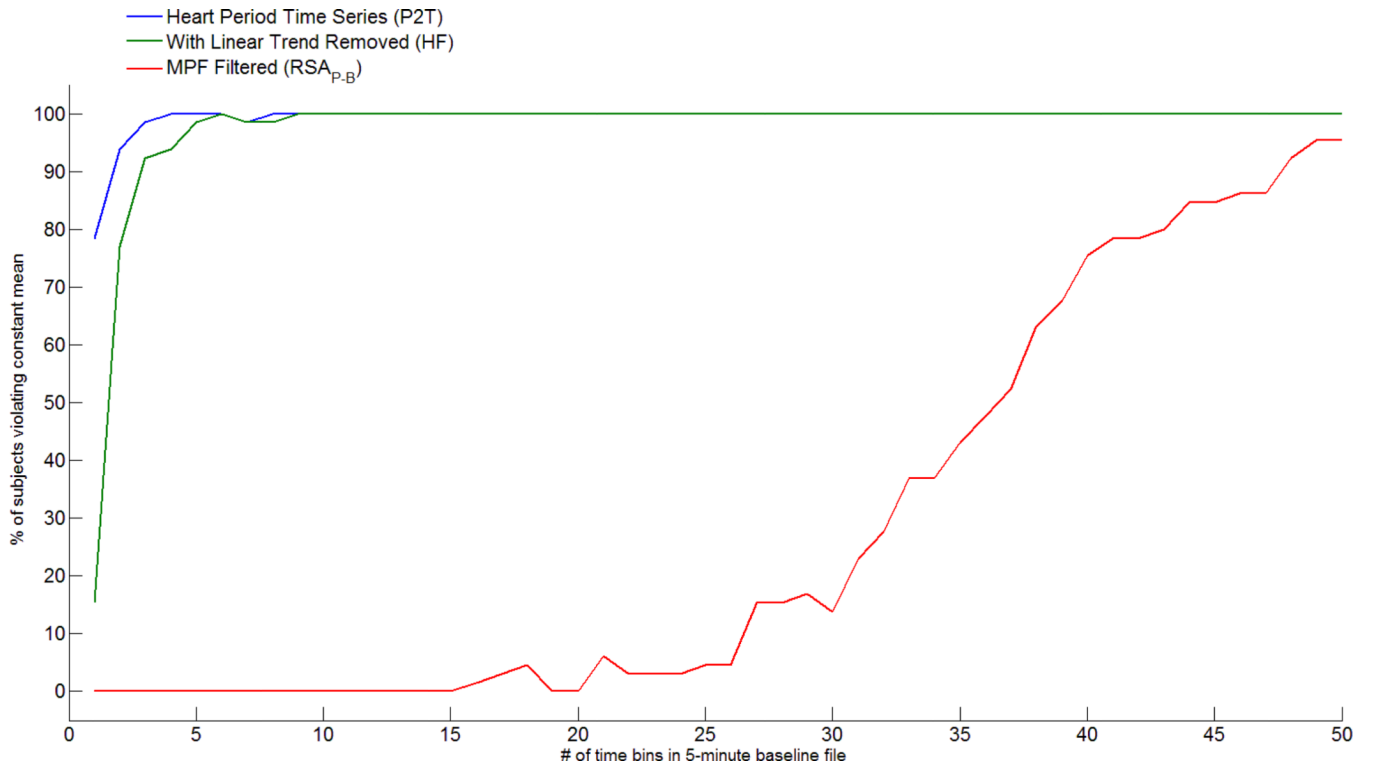
## References

- Aiken, LS.; West, SG. Multiple Regression: Testing and interpreting interactions. Newbury Park, CA: 1991.
- Allen JJB, Chambers AS, Towers DN. The many metrics of cardiac chronotropy: A pragmatic primer and a brief comparison of metrics. *Biological Psychology*. 2007; 74:243–262. [PubMed: 17070982]
- Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. *The Statistician*. 1983; 32:307–317.
- Anderson, TW. Introduction to multivariate statistical analysis. New York: John Wiley & Sons, Inc; 1958.
- Berntson GG, Bigger JT, Eckberg DL, Grossman P, Kaufmann PG, Malik M, Nagaraja HN, Porges SW, Saul JP, Stone PH, van der Molen MW. Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology*. 1997; 34:623–648. [PubMed: 9401419]
- Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; i:307–310. [PubMed: 2868172]
- Bohrer, RE.; Porges, SW. The application of time-series statistics to psychological research: An introduction. In: Keren, G., editor. *Psychological Statistics*. Hillsdale, NJ: Erlbaum; 1982. p. 309-345.

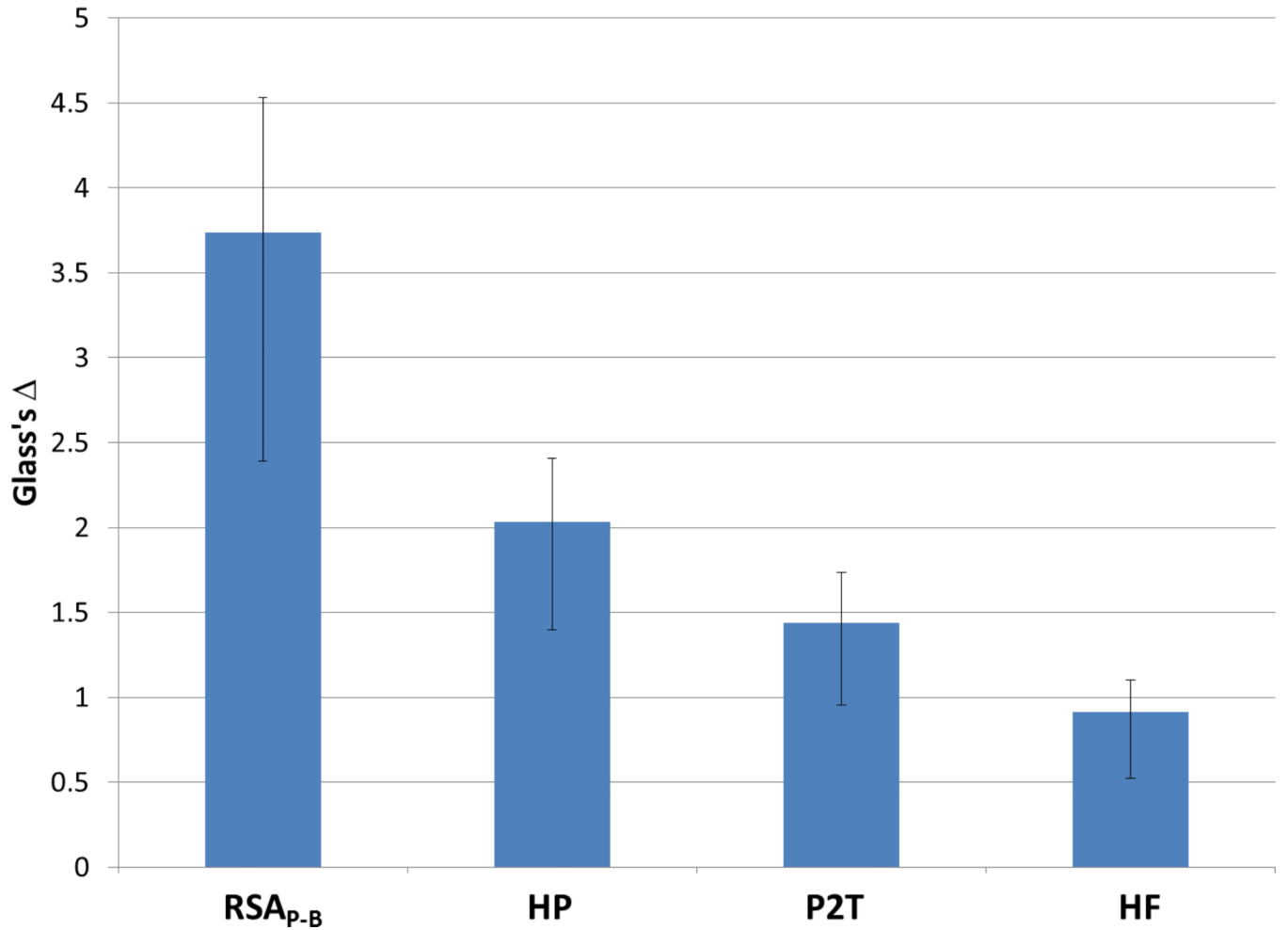
- Box GEP. A general distribution theory for a class of likelihood criteria. *Biometrika*. 1949; 36:317–346. [PubMed: 15402070]
- Box GEP. Non-normality and tests on variance. *Biometrika*. 1953; 40(3–4):318–335.
- Byrne EA, Porges SW. Data-dependent filter characteristics of peak-valley respiratory sinus arrhythmia estimation: A cautionary note. *Psychophysiology*. 1993; 30:397–404. [PubMed: 8327625]
- Camm AJ, Malik M, Bigger JT, Breithardt G, Cerutt IS, Cohen RJ, et al. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *European Heart Journal*. 1996; 17(3):354–381. [PubMed: 8737210]
- Cheng Z, Powley TL. Nucleus ambiguus projections to cardiac ganglia of rat atria: An anterograde tracing study. *Journal of Comparative Neurology*. 2000; 424:588–606. [PubMed: 10931483]
- Cohen, J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press; 1977. Rev. ed.
- Denver JW, Reed SF, Porges SW. Methodological issues in the quantification of respiratory sinus arrhythmia. *Biological Psychology*. 2007; 74:286–294. [PubMed: 17067734]
- Fouad FM, Tarazi RC, Ferrario CM, Fighaly S, Alicandri C. Assessment of parasympathetic control of heart rate by a noninvasive method. *American Journal of Physiology - Heart and Circulatory Physiology*. 1984; 246:H838–H842.
- Glass GV. Primary, secondary, and meta-analysis of research. *Educational Researcher*. 1976; 5:3–8.
- Glass, GV.; McGaw, B.; Smith, ML. *Meta-analysis in Social Research*. Beverly Hills, CA: SAGE Publications; 1981.
- Goedhart AD, Van Der Sluis S, Houtveen JH, Willemsen G, De Geus EJC. Comparison of time and frequency domain measures of RSA in ambulatory recordings. *Psychophysiology*. 2007; 44:203–215. [PubMed: 17343704]
- Grossman P, Taylor EW. Toward understanding respiratory sinus arrhythmia: relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*. 2007; 74(2):263–285. [PubMed: 17081672]
- Grossman P. Breathing Rhythms of the Heart in a World of No Steady State: A Comment on Weber, Molenaar, and van der Molen. *Psychophysiology*. 1992; 29(1):66–72. [PubMed: 1609028]
- Grossman P, Brinkman A, De Vries J. Cardiac Autonomic Mechanisms Associated with Borderline Hypertension under Varying Behavioral Demands: Evidence for Attenuated Parasympathetic Tone but Not for Enhanced Beta-Adrenergic Activity. *Psychophysiology*. 1992; 29(6):698–711. [PubMed: 1334272]
- Grossman P, Karemaker J, Wieling W. Prediction of tonic parasympathetic cardiac control using respiratory sinus arrhythmia: The need for respiratory control. *Psychophysiology*. 1991; 28:201–216. [PubMed: 1946886]
- Grossman P, Svebak S. Respiratory Sinus Arrhythmia as an Index of Parasympathetic Cardiac Control During Active Coping. *Psychophysiology*. 1987; 24(2):228–235. [PubMed: 3602275]
- Grossman P, Kollai M. Respiratory sinus arrhythmia, cardiac vagal tone, and respiration: Within- and between-individual relations. *Psychophysiology*. 1993; 30(5):486–495. [PubMed: 8416075]
- Grossman P, van Beek J, Wientjes C. A Comparison of Three Quantification Methods for Estimation of Respiratory Sinus Arrhythmia. *Psychophysiology*. 1990; 27(6):702–714. [PubMed: 2100356]
- Grossman P, Wilhelm FH, Spoerle M. Respiratory sinus arrhythmia, cardiac vagal control, and daily activity. *American Journal of Physiology: Heart and Circulatory Physiology*. 2004; 287:H728–H734. [PubMed: 14751862]
- Hayes AF, Matthes J. Computational procedures for probing interactions in linear and logistic regression: SPSS and SAS implementations. *Behavior Research Methods*. 2009; 41:924–936. [PubMed: 19587209]
- Heilman KJ, Porges SW. Accuracy of the LifeShirt (Vivometrics) in the detection of cardiac rhythms. *Biological Psychology*. 2007; 75(3):300–305. [PubMed: 17540493]
- Hering HE. A functional test of heart vagi in man. *Munch Med Wochenschr*. 1910; 57:1930.

- Hirsch JA, Bishop B. Respiratory sinus arrhythmia in humans: How breathing pattern modulates heart rate. *American Journal of Physiology*. 1981; 241:H620–H629. [PubMed: 7315987]
- Holmbeck GN. Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. *Journal of pediatric psychology*. 2002; 27(1):87–96. [PubMed: 11726683]
- Katona PG, Jih F. Respiratory sinus arrhythmia: A noninvasive measure of parasympathetic cardiac control. *Journal of Applied Physiology*. 1975; 39:801–805. [PubMed: 1184518]
- Kraemer HC, Wilson GT, Fairburn CG, Agras WS. Mediators and moderators of treatment effects in randomized clinical trials. *Archives of General Psychiatry*. 2002; 59:877–883. [PubMed: 12365874]
- Low, PA.; Sletten, DM. Laboratory Evaluation of Autonomic Failure. In: Low, PA., editor. *Clinical autonomic disorders: Evaluation and management*. Philadelphia: Lippincott-Raven; 1997. p. 130-163.
- Manly, BFJ. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. London: Chapman & Hall; 1997.
- Montano M, Ruscone TG, Porta A, Lombardi F, Pagani M, and Malliani A. Power spectrum analysis of heart rate variability to assess the changes in sympathovagal balance during graded orthostatic tilt. *Circulation*. 1994; 90:1826–1831. [PubMed: 7923668]
- Porges SW, Byrne EA. Research methods for measurement of heart rate and respiration. *Biological Psychology*. 1992; 34:193–130.
- Porges, SW., inventor. Method and apparatus for evaluating rhythmic oscillations in aperiodic physiological response systems. US Patent. 4510944. 1985 April 16.
- Porges, SW. Respiratory sinus arrhythmia: Physiological basis, quantitative methods, and clinical implications. In: Grossman, P.; Janssen, K.; Vaitl, D., editors. *Cardiorespiratory and cardiosomatic psychophysiology*. New York: Plenum; 1986. p. 101-115.
- Porges SW. The Polyvagal Perspective. *Biological Psychology*. 2007 February; volume 74(Issue 2): 116–143. [PubMed: 17049418]
- Porges, SW.; Bohrer, RE. Analyses of periodic processes in psychophysiological research. In: Cacioppo, JT.; Tassinary, LG., editors. *Principles of Psychophysiology: Physical, Social, and Inferential Elements*. New York: Cambridge University Press; 1990. p. 708-753.
- Rentero N, Cividjian A, Trevaks D, Pequignot JM, Quintin L, McAllen RM. Activity patterns of cardiac vagal motoneurons in rat nucleus ambiguus. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology*. 2002; 283:R1327–R1334.
- Riniolo T, Porges SW. Evaluating group distributional characteristics: Why psychophysicologists should be interested in qualitative departures from the normal distribution. *Psychophysiology*. 2000; 37:21–28. [PubMed: 10705764]
- Ritz T. Studying noninvasive indices of vagal control: the need for respiratory control and the problem of target specificity. *Biological Psychology*. 2009; 80(2):158–168. [PubMed: 18775468]
- Ritz T, Thöns M, Fahrenkrug S, Dahme B. The airways, respiration, and respiratory sinus arrhythmia during picture viewing. *Psychophysiology*. 2005; 42(5):568–578. [PubMed: 16176379]
- Seber, GAF. *Multivariate observations*. New York: John Wiley & Sons, Inc; 1984.
- Uijtdehaage SHJ. A BASIC program for the peak-to-valley estimation of respiratory sinus arrhythmia. *International Journal of Bio-Medical Computing*. 1994; 35:169–192. [PubMed: 8005711]
- Wilhelm FH, Grossman P, Coyle MA. Improving estimation of cardiac vagal tone during spontaneous breathing using a paced breathing calibration. *Biomedical Science Instrumentation*. 2004; 40:317–324.

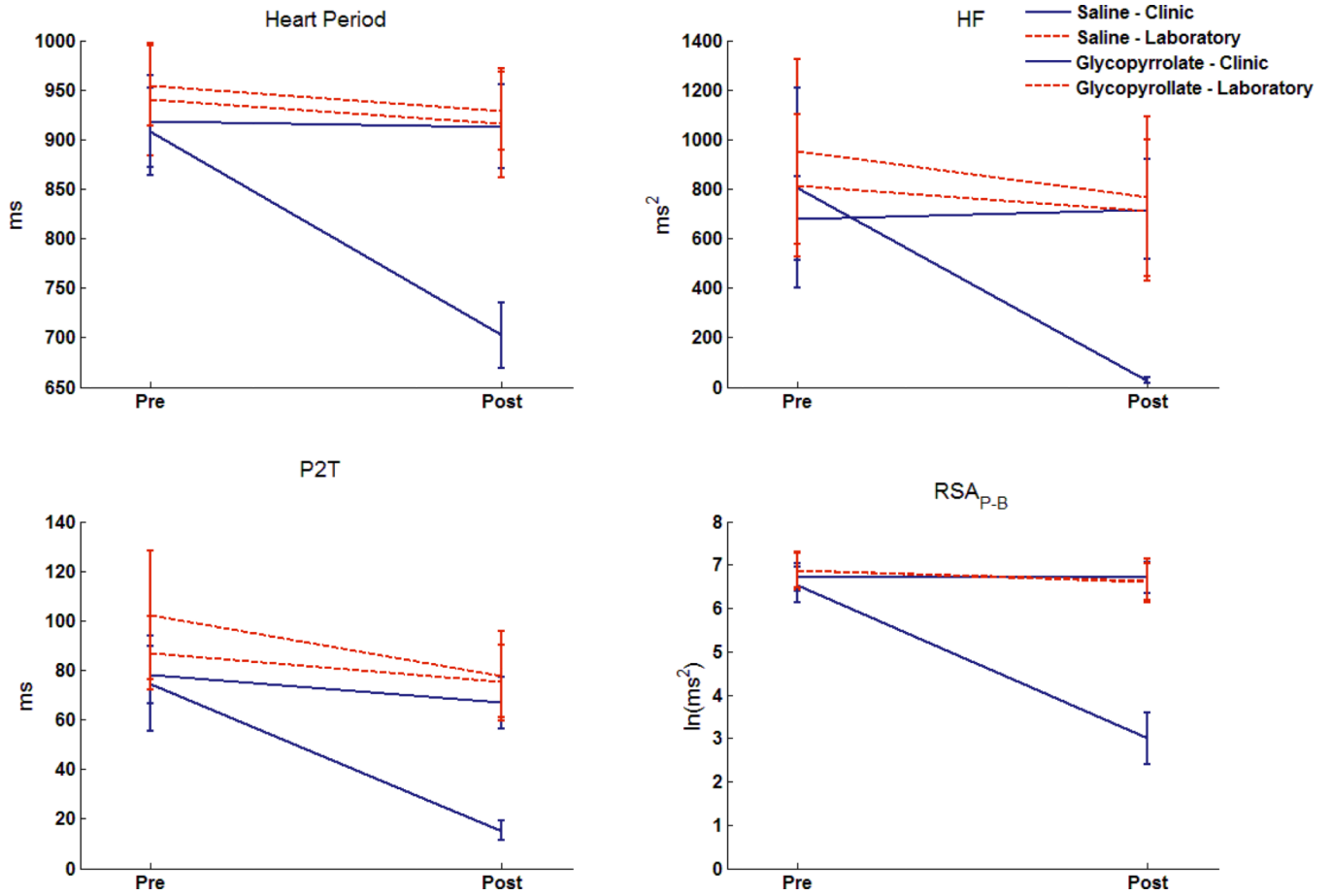




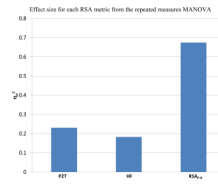
**Figure 1.** Violations of stationary mean in the heart period time series.

Bootstrap estimates of Glass's  $\Delta$ 

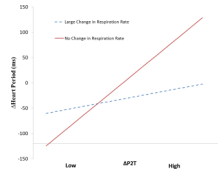
**Figure 2.** Bootstrap estimates of Glass's  $\Delta$ . Error bars denote the 95% confidence interval for Glass's  $\Delta$ . Glass's  $\Delta$  greater than 0.8 is interpreted as large (Cohen, 1977).  $N = 25$ .



**Figure 3.** Distribution of each measure in the repeated measures MANOVA: Location(2)×Time(2) ×Infusion(2). Error bars denote  $\pm 2$  standard errors of the mean. N=47.

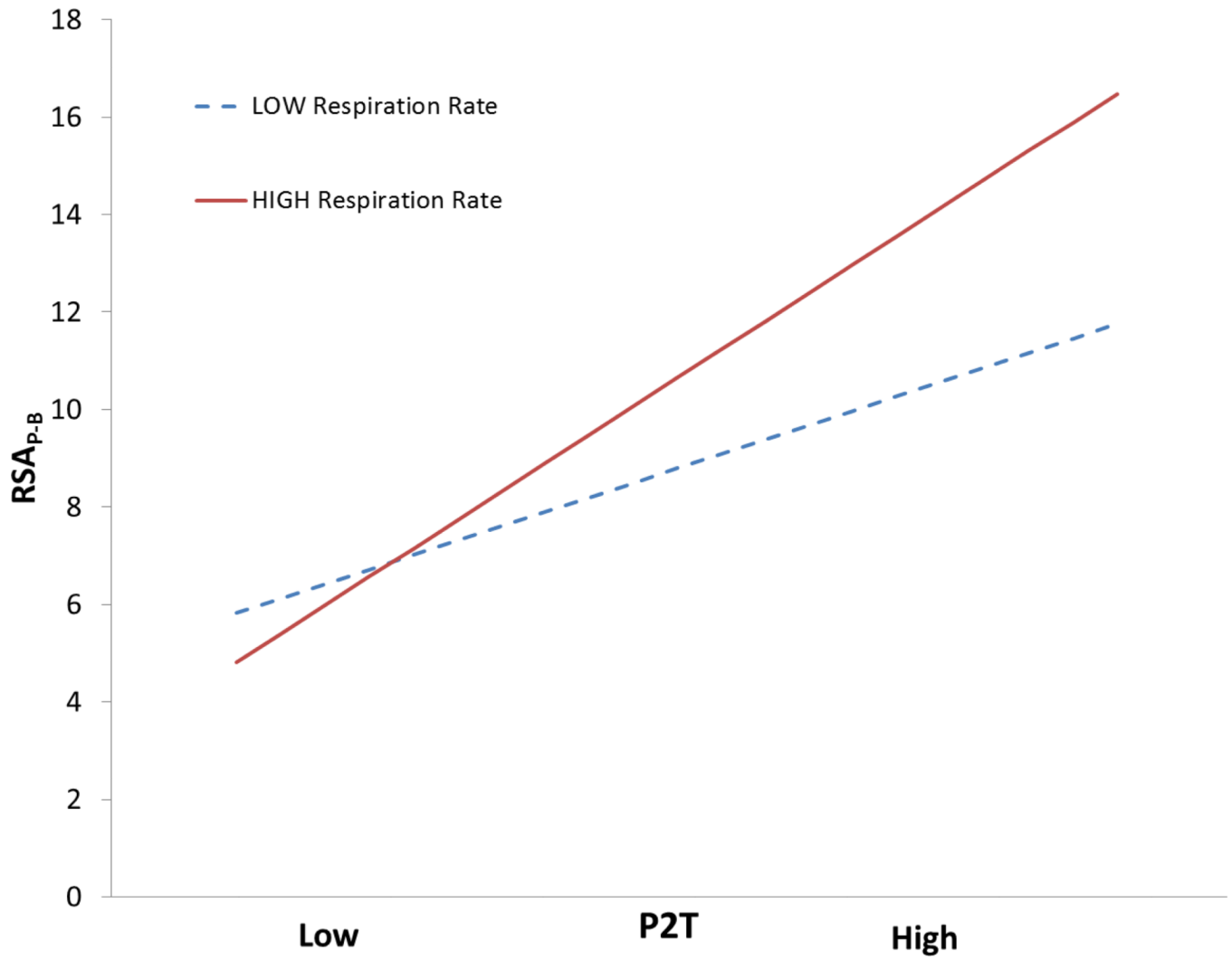


**Figure 4.** Effect size of the repeated measures MANOVA for each metric. N = 47/

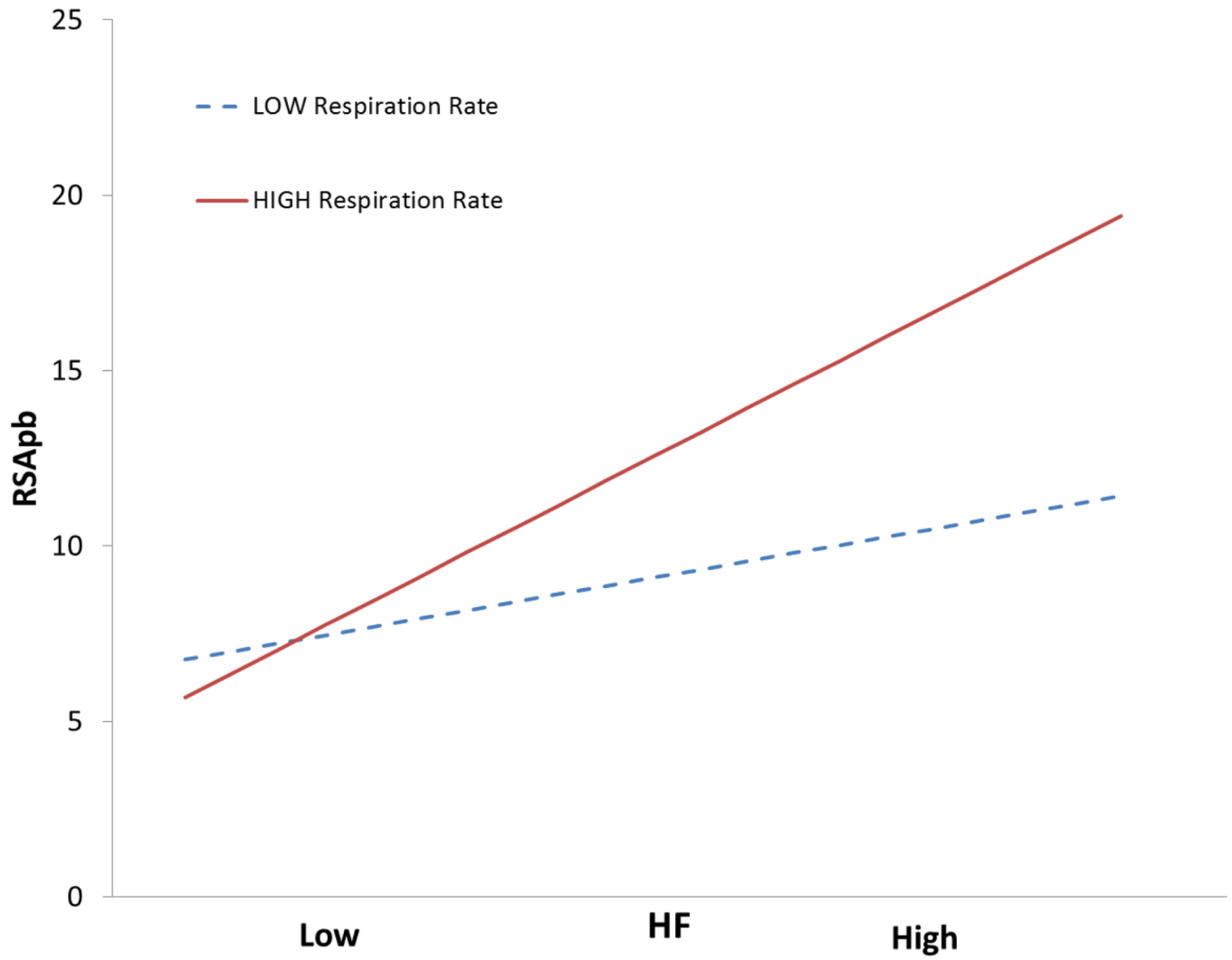


**Figure 5.**

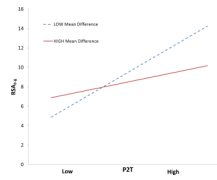
Simple slopes of  $\Delta P2T$  on  $\Delta$ Heart Period at high and zero change in respiration rate following saline infusion. The high change group represents a sample drawn at +1 SD of the absolute change in respiration rate, 0.067 Hz. The zero change group is drawn at exactly 0 change. -1 SD is exactly 0.00085 Hz. Slopes calculated with software designed by Hayes and Matthes (2009).  $N = 24$ .



**Figure 6.** Simple slopes of P2T on  $RSA_{P-B}$  at high and low respiration rates. Slopes calculated with software designed by Hayes and Matthes (2009).

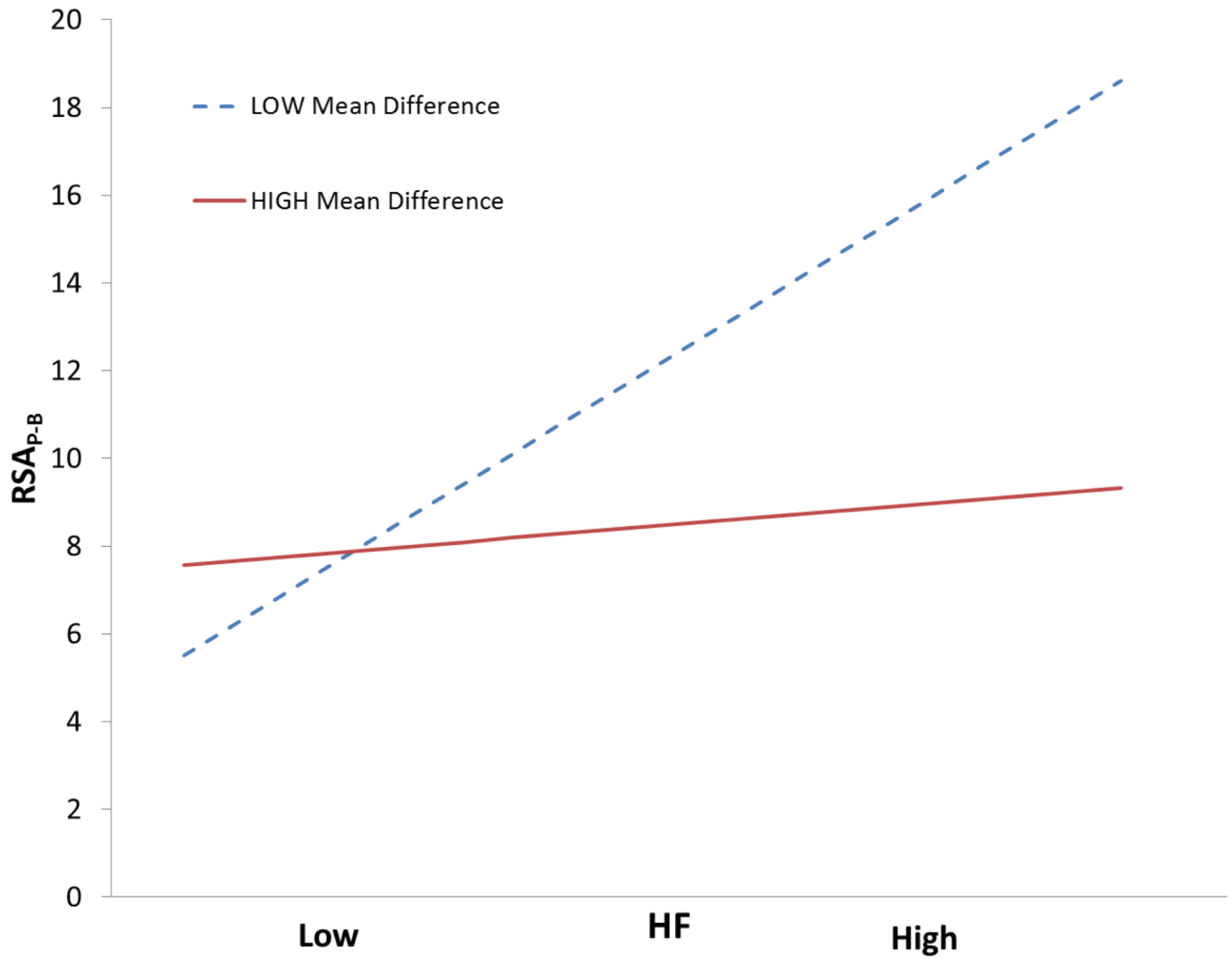


**Figure 7.** Simple slopes of P2T on  $RSAP_{P-B}$  at high and low tidal volumes. Slopes calculated with software designed by Hayes and Matthes (2009).

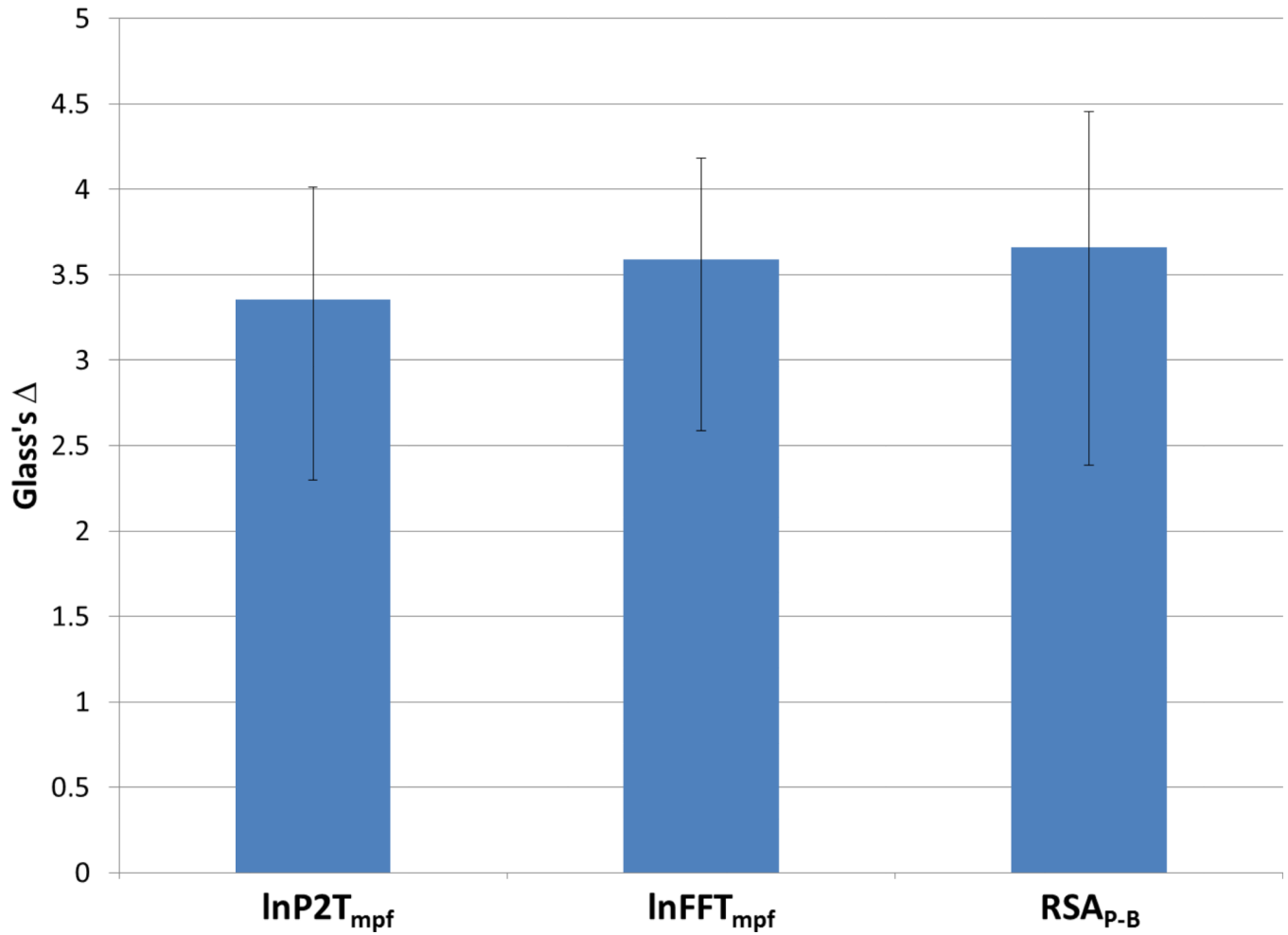


**Figure 8.** Simple slopes of P2T on  $RSA_{p-B}$  at high and low mean difference among 15 bins in the heart period time series. Slopes calculated with software designed by Hayes and Matthes (2009).





**Figure 9.** Simple slopes of HF on  $RSA_{P-B}$  at high and low mean difference among 15 bins in the heart period time series. Slopes calculated with software designed by Hayes and Matthes (2009).

Bootstrap estimates of Glass's  $\Delta$ 

**Figure 10.** Bootstrap estimates of Glass's  $\Delta$ . Error bars denote the 95% confidence interval for Glass's  $\Delta$ . Glass's  $\Delta$  greater than 0.8 is interpreted as large (Cohen, 1977).  $N = 25$ .

**Table 1**

Between subject correlations of RSA metrics

Laboratory	P2T		HF		Clinic	P2T		HF	
	RSA-P-B	P2T	RSA-P-B	HF		RSA-P-B	HF	RSA-P-B	HF
P2T	0.83**	-	-	-	P2T	0.80**	-	-	-
HF	0.71**	0.92**	-	-	HF	0.76**	0.79**	-	-
Heart Period (ms)	0.55**	0.36**	0.30*	0.30*	Heart Period (ms)	0.43**	0.48**	0.25	0.25

\*\*  $p < 0.01$ ,

\*  $p < 0.05$ , N = 65 (Laboratory), N = 50 (Clinic)

**Table 2**

## Split-half correlations of RSA metrics

	<b>RSA<sub>P-B</sub></b>	<b>P2T</b>	<b>HF</b>
First Half with Total	0.97	0.98	0.99
Second Half with Total	0.99	0.98	0.98
First with Second Half	0.93	0.91	0.94

N = 65.

**Table 3**

Split segment correlations of RSAp-B

	Minutes			30 second					
	1	2	3	1	2	3	4	5	6
Epoch with Total	0.95	0.97	0.96	0.88	0.93	0.95	0.94	0.93	0.95

N = 65.

**Table 4**

## Descriptive statistics

	Skewness	Kurtosis	Mean	Standard deviation
RSAP <sub>p-B</sub> (ln(ms <sup>2</sup> ))	-0.321	0.991	6.92	1.26
P2T (ms)	2.535	9.844	99.51	68.41
HF (ms <sup>2</sup> )	4.73	27.75	1126.31	1656.10
Respiration Rate (Hz)	-0.386	1.248	0.27	0.05
Tidal volume (mL)	1.536	2.605	477.25	211.59
Heart period (ms)	-0.120	0.218	956.27	132.81
Heart rate (bpm)	1.127	2.133	64.04	9.67

N = 65.

**Table 5**

Equality of the covariance matrices with a random grouping factor (infusion) at the research laboratory

	<b>Box's M</b>	<b>F</b>	<b>p</b>
RSA_P-B	2.95	0.94	0.42
HF	19.12	6.07	0.001
P2T	17.60	5.59	0.001

N = 48.

**Table 6**

Correlations between respiratory parameters and RSA metrics

	Respiration Rate	Tidal Volume	ln(Tidal volume) <sup>1</sup>
RSA <sub>p-B</sub>	-0.30* (-.30)	0.17 (.17)	0.19 (.19)
P2T (ms)	-0.40* (-.40)	0.34* (.33)	0.34* (.33)
HF	-0.23 (-.24)	0.19 (.20)	0.20 (.21)

\*  $p < 0.05$ ,  $N = 65$ . Sample correlations. In parentheses, bootstrapped correlations based on resampling with replacement 10,000 times.

<sup>1</sup> Correlations between the RSA metrics and the natural logarithm of tidal volume were calculated, since the kurtosis of tidal volume violated the assumption of normality.



**Table 7**

Estimated sample size needed to detect a significant correlation between the parameters 50% of the time

<b>N 50% significant</b>	<b>RSA<sub>P,B</sub></b>	<b>P2T</b>	<b>HF</b>
Respiration Rate	42	27	63
Tidal Volume	122	32	104

**Table 8**

Effect of Glycopyrrolate infusion on physiological variables

	<b>Pre-Infusion Baseline</b>	<b>Post-Infusion Baseline</b>	<b>F value</b>	<b>p value</b>
Heart Period (ms)	904.03 (107.54)	694.35 (87.51)	198.09	<0.001
HF (ms <sup>2</sup> )	731.10 (723.34)	20.11 (21.45)	23.99	<0.001
P2T (ms)	75.84 (46.30)	14.84 (9.80)	42.18	<0.001
RSA <sub>P-B</sub> (ln(ms <sup>2</sup> ))	6.56 (1.03)	2.99 (1.37)	161.02	<0.001
Respiration Rate (Hz)	0.28 (0.05)	0.29 (0.07)	1.63	0.215
Tidal Volume (mL)	501.78 (188.53)	483.66 (231.01)	0.65	0.428

N = 25.

**Table 9**

Between subject correlations between change in each RSA metric (Pre to Post saline infusion)

	$\Delta HF$	$\Delta RSA_{P-B}$
$\Delta P2T$	0.79**	0.74**
$\Delta RSA_{P-B}$	0.81**	

\*\*  
 $p < 0.01$ ,  $N = 25$ .

**Table 10**

Between subject correlations between change in heart period and change in RSA metrics (Pre to Post saline infusion)

	<u><math>\Delta</math>Heart Period</u>
$\Delta$ P2T	0.35
$\Delta$ HF	0.44*
$\Delta$ RSA <sub>P-B</sub>	0.59**
$\Delta$ Respiration Rate	-0.32
$\Delta$ Tidal Volume	-0.15

\*\*  
 $p < 0.01$ ,

\*  
 $p < 0.05$ ,  $N = 25$ .

**Table 11** $\Delta$ RSA  $\times$   $\Delta$ Respiration Parameter interaction term in moderation analyses predicting change in heart period

<b>A</b>			
$\Delta$ Respiration Rate	$\Delta$ RSA <sub>p-B</sub>	$\Delta$ P2T	$\Delta$ HF
p-value	0.20	0.02	0.08
<i>t</i> (21)	-1.32	-2.61	-1.82
R <sup>2</sup> Change	0.05	0.20	0.12
Total R <sup>2</sup>	0.40	0.39	0.26

<b>B</b>			
$\Delta$ Tidal Volume	$\Delta$ RSA <sub>p-B</sub>	$\Delta$ P2T	$\Delta$ HF
p-value	0.22	0.22	0.05
<i>t</i> (21)	-1.26	-1.26	-2.08
R <sup>2</sup> Change	0.04	0.04	0.14
Total R <sup>2</sup>	0.49	0.41	0.33

N = 25.

**Table 12**RSA  $\times$  Respiration Parameter interaction term in moderation analyses predicting RSA<sub>P-B</sub>

<b>A</b>		
<b>Respiration Rate</b>	<b>P2T</b>	<b>HF</b>
p-value	<0.001	<0.001
<i>t</i> (61)	7.28	4.54
R <sup>2</sup> Change	0.14	0.12
Total R <sup>2</sup>	0.84	0.64

<b>B</b>		
<b>Tidal Volume</b>	<b>P2T</b>	<b>HF</b>
p-value	<0.001	0.02
<i>t</i> (61)	-4.29	-2.36
R <sup>2</sup> Change	0.07	0.04
Total R <sup>2</sup>	0.77	0.54

N = 65.

**Table 13**RSA  $\times$  Average heart period difference (15 bins) interaction term in moderation analyses predicting RSA<sub>P-B</sub>

	<b>P2T</b>	<b>HF</b>
p-value	<0.001	<0.001
<i>t</i> (61)	-7.57	-5.29
R <sup>2</sup> Change	0.10	0.09
Total R <sup>2</sup>	0.89	0.80

N = 65.

**Table 14**

## Descriptive statistics

	Skewness	Kurtosis	Mean	Standard deviation
$\text{RSA}_{\text{P-B}} (\ln(\text{ms}^2))^*$	-0.427	1.266	7.03	1.25
$\ln\text{P2T} (\ln(\text{ms}^2))$	-0.178	0.583	7.17	1.14
$\ln\text{HF} (\ln(\text{ms}^2))$	-0.226	0.471	6.45	1.10

N = 65.

\* For comparison with other RSA measures, this value was obtained without dividing the filtered heart period time series into 30 second epochs.



**Table 15**

RSA metric feature summary

	<b>RSA<sub>P-B</sub></b>	<b>P2T</b>	<b>HF</b>
Stability	High	High	High
Kurtosis	Conforms	Violates	Violates
Skewness	Conforms	Violates	Violates
Homoscedasticity	Conforms	Violates	Violates