# Effortful Listening: Sympathetic Activity Varies as a Function of Listening Demand but Parasympathetic Activity Does not

Kate Slade[a], Sophia E. Kramer[b], Stephen Fairclough[c], and Michael Richter[c]

[a] Department of Psychology, Lancaster University, Lancaster, UK
[b] Amsterdam UMC, Vrije Universiteit Amsterdam, Otolaryngology—Head and Neck Surgery, Ear and Hearing, Amsterdam Public Health Research Institute, De Boelelaan, Amsterdam, the Netherlands
[c] School of Psychology, Liverpool John Moores University, Liverpool, UK

Correspondence concerning this article should be addressed to Kate Slade, Department of Psychology, Fylde College, LA1 4YF, Lancaster, UK or Michael Richter, Faculty of Health, School of Psychology, Byrom Street, L3 3AF, Liverpool, UK. E-mail: k.slade2@lancaster.ac.uk, m.richter@ljmu.ac.uk

## Highlights

- Increased listening demand leads to increased cardiac sympathetic activity.
- Increased listening demand results in increased PEP reactivity.
- Extremely high (impossible) listening demand results in weak ANS response.

## Abstract

Research on listening effort has used various physiological measures to examine the biological correlates of listening effort but a systematic examination of the impact of listening demand on cardiac autonomic nervous system activity is still lacking. The presented study aimed to close this gap by assessing cardiac sympathetic and parasympathetic responses to variations in listening demand. For this purpose, 45 participants performed four speech-in-noise tasks differing in listening demand—manipulated as signal-to-noise ratio varying between +23 dB and -16 dB—while their pre-ejection period and respiratory sinus arrythmia responses were assessed. Cardiac responses showed the expected effect of listening demand on sympathetic activity, but failed to provide evidence for the expected listening demand impact on parasympathetic activity: Pre-ejection period reactivity increased with increasing listening demand across the three possible listening conditions and was low in the very high (impossible) demand condition, whereas respiratory sinus arrythmia did not show this pattern. These findings have two main implications. First, cardiac sympathetic responses seem to be the more sensitive correlate of the impact of task demand on listening effort compared to cardiac parasympathetic responses. Second, very high listening demand may lead to disengagement and correspondingly low effort and reduced cardiac sympathetic response.

*Keywords:* effort; sympathetic activity; parasympathetic activity; pre-ejection period; respiratory sinus arrythmia; motivational intensity theory;

## 1. Introduction

In the last decade, physiological measures have become popular in the literature and research on listening effort. Researchers used various measures like pupil dilation (Koelewijn, Zekveld, Lunner, & Kramer, 2018b; Strand, Brown, Merchant, Brown, & Smith, 2018; Zekveld & Kramer, 2014), electroencephalographic (EEG) activity (Bernarding, Strauss, Hannemann, Seidler, & Corona-Strauss, 2017; Miles et al., 2017), pre-ejection period (Plain et al., 2020; Richter, 2016a), skin conductance (Alhanbali, Dawes, Millman, & Munro, 2019; Mackersie & Calderon-Moultrie, 2016; Mackersie & Kearney, 2017; Seeman & Sims, 2015), electromyographic activity (Mackersie & Cones, 2011), heart rate variability (Mackersie & Calderon-Moultrie, 2016; Mackersie & Kearney, 2017; Seeman & Sims, 2015), and fMRI responses (Wild et al., 2012) to assess the effort that individuals invest in listening tasks (see Francis & Love, 2020; McGarrigle et al., 2014, for reviews). However, given that the selection of a particular physiological measure in listening effort research was frequently unaccompanied by a theoretical rationale, the current psychophysiological literature on the topic is fragmented. In this article, we draw from empirical evidence on autonomic nervous system activity associated with physical effort as well as on motivational intensity theory (Brehm & Self, 1989) to present a model that enables a more systematic approach to researching the (cardiovascular) psychophysiology of listening effort and provide a first empirical test of this model. A more systematic, theory-driven approach will help researchers to examine listening effort in a more focussed manner. It will provide guidance which measures to assess and which effects to expect. It will also facilitate the aggregation of individual studies on the psychophysiology of listening effort in systematic reviews and make these reviews more conclusive.

Motivational intensity theory (Brehm & Self, 1989) is a psychological theory about effort investment that adopts a definition of effort similar to the definition of listening effort provided by the Fifth Eriksholm Workshop on “Hearing Impairment and Cognitive Energy” (Pichora-Fuller et al., 2016): (Listening) effort refers to energy or resources that are used to overcome obstacles in goal-directed tasks (for instance, watching a movie on TV while your neighbours are having a noisy birthday party). Motivational intensity theory suggests that these resources are limited and that individuals therefore aim to conserve them whenever possible. Consequently, individuals use available information about task demand—that is, information about the amount of resources required to successfully perform the task at hand—to adjust their effort investment: the lower the demand, the lower the effort investment. This strategy ensures that individuals never waste resources by investing more than necessary. However, the proportional relationship between task demand and effort investment requires an upper limit to avoid wasting resources by investing more effort than justified or by investing effort when task demand becomes so high that success is impossible. Consequently, motivational intensity theory predicts that task demand directly determines effort if 1) the importance of success justifies the required effort investment and if 2) task success is possible. If these two conditions are not met, individuals should refrain from investing effort (see Richter, 2013; Wright, 2008, for detailed discussions of motivational intensity theory's predictions).

Most of the empirical research on motivational intensity theory has relied on Wright’s (1996) suggestion that effort investment in cognitive tasks (i.e., mental effort) is associated with increased myocardial sympathetic nervous system (SNS) activity. Drawing on this perspective, researchers examined the impact of various manipulations of task demand and success importance on cardiovascular parameters affected by sympathetic activity, like pre-ejection period and systolic blood pressure (Gendolla, Wright, & Richter, 2019; Richter, Gendolla, & Wright, 2016, for recent overviews). Given Wright’s (1996) focus on myocardial sympathetic activity, it comes as no surprise that research on motivational intensity theory has rarely examined the association between effort and the activity of the parasympathetic nervous system (PNS) (see Harper, Eddington, & Silvia, 2016; Richter, 2010b; Silvia, Beaty, Nusbaum, Eddington, & Kwapil, 2014; Silvia, Eddington, Beaty, Nusbaum, & Kwapil, 2013; Silvia et al., 2016; Venables & Fairclough, 2009, for exceptions).

Interestingly, the physiological literature on physical effort suggests that both branches of the autonomic nervous system (ANS) are involved in effortful tasks (McArdle, Katch, & Katch, 2010;

Michael, Graham, & Davis, 2017). The increase in cardiac activity that accompanies physical exercise is the result of both decreased PNS activity and increased SNS activity. The relative contribution of the two systems differs however as a function of the intensity of the physical exercise (Robinson, Epstein, Beiser, & Braunwald, 1966; White & Raven, 2014). The increase in cardiac activity from rest to low-intensity physical exercise is mainly driven by reductions in inhibiting PNS activity. The contribution of the SNS is negligible. However, both the PNS and the SNS contribute to the additional increase in cardiac activity from low-intensity exercise to moderate-intensity exercise: PNS activity decreases further and SNS activity increases. Given that PNS withdrawal is almost complete at moderate exercise intensity levels, increases in cardiac activity from moderate to high levels of exercise intensity are mainly driven by additional increases in SNS activity. Increases in physical effort—from low to high intensity exercise—are thus characterised by a change from an uncoupled parasympathetic withdrawal mode of autonomic control to a coupled reciprocal mode (Berntson, Cacioppo, & Quigley, 1991) and by a specific change in SNS-PNS balance: PNS activity dominates if physical effort is low whereas SNS activity dominates if physical effort is high.

Drawing on models where patterns of ANS activity during performance of demanding (stressful) cognitive tasks are hypothesised to reflect adaptive physiological responses to physical threats in ancestral environments (Boyce & Ellis, 2005; Nesse, Bhatnagar, & Ellis, 2016; Nesse, Bhatnagar, & Young, 2007; Obrist, 1981), we suggest that our ANS system does not differentiate between physical and cognitive demands in relation to their impact on the heart. Consequently, the same autonomic mechanisms associated with physical effort should underlie effort investment in cognitive tasks—including tasks that require the investment of listening effort. Therefore, low mental (listening) effort should be associated with decreased PNS activity and negligible increases in SNS activity. Moderate mental (listening) effort should be characterised by strong reductions in PNS activity and increased SNS activity. High mental (listening) effort should be associated with complete PNS withdrawal and strong increases in SNS activity. Figure 1 illustrates this pattern modelled as quadratic relationships between effort intensity and SNS and PNS activity. Appendix A provides information on why we decided to use quadratic functions to model the relationship between effort intensity and SNS and PNS activity.

The existing empirical literature on motivational intensity theory and listening effort does not provide conclusive evidence regarding this hypothesis. Studies on motivational intensity theory that included measures of both SNS and PNS activity had complex designs that make a straightforward interpretation difficult. The studies examined the impact of perfectionism (Harper et al., 2016), grit (Silvia et al., 2013), creativity (Silvia, Beaty, et al., 2014), dysphoria (Silvia et al., 2016; Silvia, Nusbaum, Eddington, Beaty, & Kwapil, 2014), reward value (Richter, 2010b), task context (Richter, 2010b), and bogus performance feedback (Venables & Fairclough, 2009) but did not—with one exception (Silvia et al., 2016)—include direct manipulations of task demand, which provide the most straightforward test of the predicted relationship between ANS activity and mental effort. Silvia and colleagues (2016) examined the interaction of task difficulty and depression in a d2 concentration task—a task in which one has to find all “d’s” with two dashes in a series of letters presented with up to four dashes (Brickenkamp, 2002). They observed that SNS activity—assessed as pre-ejection period reactivity—increased with increasing task difficulty from the easy condition to the hard condition but was low in the very-hard condition. However, only participants with a high number of depressive symptoms displayed this pattern. If participants’ depression levels were low, task difficulty did not affect SNS activity. Moreover, Silvia and colleagues did not observe any effects on PNS-assessed as respiratory sinus arrhythmia reactivity-activity. The absence of effects on PNS activity is characteristic for most studies on motivational intensity theory that examined PNS responses. There are, however, two exceptions. Silvia et al. (2013) and Silvia, Beaty, et al. (2014) found that both SNS and PNS activity increased from baseline to task performance. The SNS effects observed in these studies were thus in line with our predictions but the observed increases in PNS activity are difficult to interpret in terms of effort investment.

Four listening effort studies assessed SNS and PNS activity, so far. Seeman and Sims (2015) assessed changes in skin conductance—an indicator of sympathetic activity (Dawson, Schell, & Filion, 2017)—and heart rate variability—an indicator of parasympathetic activity (Berntson et al., 1997; Task Force of the European Society of Cardiology and the North American Society of Pacing Electrophysiology, 1996)—in response to two different listening tasks. In a diotic-dichotic listening task (Study 1), increases in task complexity increased heart rate variability but did not influence skin conductance level. In a speech-in-noise task (Study 2), lower signal-to-noise (SNR) ratios were associated with increased heart rate variability but no effects of SNR on skin conductance level were observed. Mackersie and Calderon-Moultrie (2016) also assessed skin conductance level and heart rate variability in a speech-in-noise task. They observed that the listening task resulted in increased skin conductance level and decreased heart rate variability compared to rest. Moreover, both measures differentiated between normal and fast speaking rates. If speaking rate was fast (i.e. if more effort was required to understand the speech), heart rate variability was lower and skin conductance level was higher than if speaking rate was normal. Mackersie and Kearney (2017) used a speech-in-noise task that included a manipulation of task demand—that is, participants had either to repeat words from spoken text (low task demand) or answer comprehension questions about the text (high task demand) —as well as a manipulation of evaluative observation—that is, participants were either recorded for later assessment or not. They found decreased heart rate variability and increased skin conductance when task demand increased. However, heart rate variability did not vary as a function of task demand or observation. Skin conductance increased in the high-demand-high-evaluation condition compared to the other three conditions. In short, the available listening effort studies that examined the activity of both ANS branches provided some support for the notion that listening effort is associated with changes to sympathetic and parasympathetic activity assessed using skin conductance, and heart rate variability, respectively. However, a consistent relationship between either branch of the ANS and listening demand was not observed: Skin conductance and heart rate variability varied as a function of listening demand in some studies, but not in others.

To gather more conclusive information about the role of sympathetic and parasympathetic activity in listening effort, we decided to examine ANS activity across multiple levels of listening demand. Manipulating listening demand across more than two levels allowed us to examine the effect of changes in listening demand on effort-related ANS activity in a more comprehensive manner. In particular, it allowed us to specifically test the predicted quadratic relationships between listening demand and SNS and PNS activity, which is not possible with only two demand levels. We also decided to include a condition with extremely high listening demand to test for the disengagement that motivational intensity theory predicts for impossible demand levels. We focussed on cardiac ANS activity given our physiological rationale and given that ANS responses show regional differentiation (Esler et al., 1990). In contrast to preceding work on listening effort, we therefore did not use skin conductance as an indicator of sympathetic activity, but pre-ejection period (PEP)—the time interval between the excitation of the left heart ventricle and the beginning of the ejection of blood into the aorta. Skin conductance level is influenced by sympathetic outflow to the sweat glands (Dawson et al., 2017), whereas PEP constitutes an indicator of SNS impact on the heart (Newlin & Levenson, 1979; Sherwood et al., 1990). To assess PNS activity we used—like preceding work on listening effort and motivational intensity theory—a specific type of heart rate variability, respiratory sinus arrythmia (RSA). RSA represents variability in the heart beat synchronous with respiratory activity and is considered a valid indicator of cardiac PNS activity (Berntson et al., 1997; Task Force of the European Society of Cardiology and the North American Society of Pacing Electrophysiology, 1996). Assessing PEP and RSA thus allowed us to specifically observe the cardiac SNS and PNS responses to variations in listening demand.

To examine how variations in listening demand affect PEP and RSA, participants performed a listening task in which they had to understand speech embedded in background noise, which was varied to create three possible and one impossible listening demand levels. We expected a quadratic increase of PEP reactivity—the change from rest to task performance—across the three possible demand levels: The relative increase in PEP reactivity from low demand to medium demand should be smaller than the increase from medium to high demand. RSA reactivity was hypothesised to show a quadratic decrease across these demand levels: The relative increase in RSA reactivity from low demand to medium demand should be greater than the increase from medium to high demand. In the impossible demand condition, we expected participants to disengage and thus predicted correspondingly low PEP and RSA reactivity. Figure 2 displays these hypotheses. Please note that for both measures a greater reactivity implies a more negative value given that increased SNS and decreased PNS activity lead to shorter PEP and RSA values.

# 2. Material and Methods

## 2.1 Participants and Design

A sample of 45 adults ($M_{\text{age}} = 24.87$, $SD_{\text{age}} = 5.74$; $M_{\text{BMI}} = 25.12$, $SD_{\text{BMI}} = 5.76$), 26 females and 19 males, without pacemakers participated for a potential 20-GBP in Amazon vouchers. Sample size was determined using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) setting alpha error to 5%, beta error to 5% and Cohen's $f$ to 0.25. All participants reported no diagnosis of hearing impairment. Each participant participated individually and completed all four demand conditions (low, moderate, high, and impossible) of a speech-in-noise task presented in random order.

## 2.2 Materials

All materials were presented to participants on a single computer screen using experiment generation software (Inquisit by Millisecond Software, Seattle, WA). The software presented all task stimuli and collected all the participants responses.

### *2.2.1 Speech-in-noise task*

In each trial of the speech-in-noise task, participants listened via headphones to a 32-second short story spoken by a female voice in the presence of white noise—the story started a few milliseconds after the white noise. Please see Appendix B for three examples of such stories, but all story stories, audio files, and associated experimental scripts can be found in the supplementary materials (https://doi.org/10.24377/LJMU.d.00000087). All stories were created by the authors using computer-generated speech in a female voice without accent. The decibel (dB) level of the white noise was informed by a pilot calibration procedure. Six individuals with normal hearing were presented with the short stories in differing levels of white noise (SNR levels from -10 dB to 2 dB).

The individuals indicated whether they found it easy, moderately difficult, difficult, or impossible to identify the speech at these SNR levels. Trials with SNRs of 2 dB and -4 dB were most frequently rated as moderately difficult and difficult, respectively, these levels were thus selected for use in the moderate and high-demand listening task conditions. However, when selecting the SNR values for the low and impossible conditions, it was necessary to ensure that the low demand condition would be sufficiently easy, and that task success would be unattainable in the impossible condition. To ensure that the low-demand condition would be sufficiently easy, an SNR level higher than those employed during piloting was chosen to remove any ambiguity in ensuring that minimal-to-no effort would be required for task success. Similarly, the SNR level selected in the impossible demand condition reflected a SNR level much lower than presented during piloting to ensure that task success would be unattainable. We decided to use this calibration procedure—and against using four SNR levels with equal SNR increases from one demand level to the next one—because research on motivational intensity theory suggested that it is the subjective perception of task demand that counts (e.g., Gendolla & Krusken, 2001; Wright, 1998; Wright & Franklin, 2004). The resulting SNR levels were 23dB in the low-demand condition, 2dB in the moderate-demand condition, -4dB in high-demand condition, and -16dB higher than the speech in the impossible-demand condition. The output volume of the experimental computer was adjusted and maintained at a volume that was not too adverse. This output level was measured with a sound level meter to ensure that the dB SPL did not exceed 80 during the experiment, and participants were asked to confirm that the volume was not too high. At the end of each short story, participants were given five seconds to respond to a 3-option multiple-choice comprehension question. The speech-in-noise task trials were presented in blocks of ten trials of one and the same demand level. The total duration of a trial was kept constant at 38.50 seconds by adapting the inter-trial break as a function of participant's response time to the multiple-choice question. Total duration of a block of the speech-in-noise task was thus 385 seconds for all participants and in all demand conditions.

### ***2.2.2 Fatigue, demand, and effort measures***

Participants' fatigue was assessed at the start of the experiment and after each block (please see Section 2.3 for details of the experimental procedure) to examine whether increases in listening effort would result in increased fatigue. A positive relationship between listening effort and fatigue has been frequently reported in the literature (Alhanbali, Dawes, Lloyd, & Munro, 2017; Hornsby, 2013) and we attempted to replicate this relationship in our specific task context. Fatigue was measured using a computer-based 9-item questionnaire designed for the purpose of this study, but items included were based on key words in existing measures (Alhanbali et al., 2017; Nachtegaal et al., 2009). Each item was composed of one fatigue-related word (fatigued, tired, and worn out) and one word referring to an alert, energised state (energised, lively, well-rested), and participants had to decide for each item which one of the two words best described their current state. We had originally planned to present all possible combinations of the terms, but due to a coding mistake the fatigue questionnaire included sometimes 10 items and up to two pairs were presented twice. To take this issue into account, we quantified self-reported fatigue as the percentage of items in which a participant selected the fatigue-related term. Participants reported perceived demand and effort after each block of 10 speech-in-noise trials using two items ("How mentally demanding was the listening task?", "How hard did you have to work to accomplish your level of performance?") adapted from the NASA Task Load Index (Hart & Staveland, 1988). The item scales ranged from 1 (*very low*) to 5 (*very high*).

### *2.2.3 Physiological measures*

For the quantification of PEP as indicator of SNS activity and RSA as indicator of PNS activity, a CardioScreen 1000 impedance cardiograph (Medis, Illmenau, Germany) collected an impedance cardiogram (ICG) and an electrocardiogram (ECG) at a sampling rate of 1000 Hz. The four pairs of disposable electrodes of the device were placed on the left and right sides of the participant's chest at the height of the xiphoid and on the right and left sides of the neck. To enable comparison with preceding work on motivational intensity theory, which has frequently used blood pressure to test effort-related hypotheses (see Gendolla et al., 2019, for a recent review), a Dinamap Carescape V100 monitor (GE Healthcare, Buckinghamshire, UK) assessed participants' systolic (SBP) and diastolic blood pressure (DBP) in two-minute intervals using the oscillometric method. The monitor's blood pressure cuff was applied to the participant's upper left arm. The collected ECG was also used to determine participants' heart rate (HR), which allowed us in combination with participants' DBP values to verify that PEP responses reflected myocardial sympathetic activity, and not pre-load or after-load effects (Obrist, 1981; Obrist, Light, James, & Strogatz, 1987; Sherwood et al., 1990).

### 2.3 Procedure

Experiment generation software (Inquisit by Millisecond Software, Seattle, WA) controlled the presentation of the experimental stimuli and collected participants' responses. After participants had provided informed consent, the experimenter (the first author) measured their height and weight. The experimenter then attached the CardioScreen electrodes and the blood pressure cuff while participants indicated their age and gender. Participants completed the fatigue measure for the first time to determine baseline fatigue.

Participants then performed the four demand versions of the speech-in-noise task in four blocks. The order of the blocks was determined by computer-controlled simple randomization. Each block included task instructions, two practice trials, a baseline period, ten speech-in-noise task trials, and the fatigue, demand, and effort items presented in the order described in the following sentences. The task instructions provided general information about the task and informed participants that they would earn an £5 Amazon Voucher if they answered correctly at least seven of the multiple-choice questions of the current block. The practice trials were of the same demand level as the ten speech-in-noise task trials presented in the block and allowed the acquisition of information about task demand. Participants received feedback on the accuracy of their response to the multiple-choice question at the end of each practice trial, but not during the main speech-in-noise task. During the 6-minute baseline period, participants watched a clip from the nature documentary Kingdom of Plants (Williams, 2012), while their cardiovascular activity at rest was assessed. ECG and ICG signals were continuously assessed during the baseline period and during the presentation of the ten speech-in-noise task trials. Blood pressure values were taken in two-minute intervals starting after 60 seconds after the beginning of baseline period and 10 seconds after the beginning of the task period. After the task, participants used the fatigue, effort, and demand items to reports their current fatigue and how effortful and demanding the preceding task block had been. After a participant had completed all four task blocks, the researcher carefully debriefed and remunerated them.

### 2.4 Data Preprocessing

The collected ICG and ECG signals were analysed offline using BlueBox software (Richter, 2010a). ECG R-peaks were automatically detected using a peak threshold detection algorithm and the detected R-peaks were visually confirmed. Ectopic beats were deleted as recommended by Lippman, Stein, and Lerman (1994). HR was then determined by counting the number of R-peaks (beats) per minute. The first derivative of the ICG signal (dZ) was computed and individual heart cycles were extracted from the resulting dZ/dt signal using the locations of the detected R-peaks. The dZ/dt segments were then averaged to obtain one ensemble average per minute (Kelsey & Guethlein, 1990). Two independent raters identified in each ensemble average R-onset and B-point following the official guidelines of the Society for Psychophysiological Research (Sherwood et al., 1990). PEP values were computed as difference between R-onset and B-point for each ensemble average and rater. The arithmetic means of the PEP values of the two raters (ICC[2, 2] > .99) constituted our final PEP scores.

Respiratory sinus arrythmia was determined following published guidelines (Berntson et al., 1997; Task Force of the European Society of Cardiology and the North American Society of Pacing Electrophysiology, 1996). The detected R-peaks were first transformed into interbeat intervals (IBIs). IBIs were resampled at 4 Hz, detrended with a 3-order polynomial (Litvack, Oberlander, Carney, & Saul, 1995), and transformed into a power spectrum by Fast Fourier Transform (Welch's method, 1024 data points, Hamming window, 50% window overlap). Following the standard approach (Berntson et al., 1997; Task Force of the European Society of Cardiology and the North American

Society of Pacing Electrophysiology, 1996), the power in the range from 0.15 to 0.40 Hz expressed in normalized units—that is, the percentage of the power in the range from 0.15 to 0.40 Hz relative to the power in the range from 0.04 to 0.40 Hz—was used as RSA measure.

Given that RSA refers to heart rate variability synchronous with respiration and that the standard range of 0.15 to 0.40 Hz might not adequately capture the specific respiration frequencies of our individual participants, we also computed a respiration-centred RSA (resp-RSA; Hernando et al., 2016; Skytioti, Sovik, & Elstad, 2017). We first determined each participant's respiration frequency using the ICG dZ signal (de Geus, Willemsen, Klaver, & van Doornen, 1995; Houtveen, Groot, & de Geus, 2006). The dZ signal was filtered with 10-Hz low-pass and 0.1-Hz high-pass Butterworth filters and then smoothed with three Savitzky-Golay filters as described in Seppa, Viik, and Hyttinen (2010).[1] The filtered signal was then downsampled to 10 Hz and transformed into a power spectrum by Fast Fourier Transform (Welch's method, 1024 data points, Hamming window, 50% window overlap). After smoothing the spectrum with a Savitzky-Golay filter (11 data points, 2nd order), the frequency associated with the spectrum's peak amplitude in the range between 0.01 and 0.50 Hz was used as the participant's respiration frequency. For ten participants the spectrum did not allow an unambiguous identification of a peak, and the resp-RSA analysis is thus based on the data of the 35 participants with a clear spectrum peak. The processing of the IBI signal followed the same procedure as for RSA except that the normalised power in the frequency band centred around the participant's respiration frequency (respiration frequency +/- 0.05 Hz) was used and that the normalisation was done in relation to the power in the band from .04 Hz to 0.50 Hz.

To obtain PEP, SBP, and DBP baseline scores, the measures obtained during the last five minutes of each baseline period were averaged. A 5-minute window was employed to allow participants the first minute during the baseline period to return to a physiologically restful state, as such the last 5 minutes were considered to best reflect the participants baseline state. PEP, SBP, and DBP task scores were computed as arithmetic mean of the measures collected during the first five

---

[1] We used a frame size of 2500 ms instead of 2000 ms for the last of the three Savitzky-Golay filters.

minutes of each task period. The first five minutes of this period were used as this was considered to be the time at which the participants would be most engaged with the task. HR, RSA, and resp-RSA values were already based on the appropriate five-minute epochs extracted from baseline and task periods. In the last step of the data preprocessing, cardiovascular reactivity (change) scores (Llabre, Spitzer, Saab, Ironson, & Schneiderman, 1991) were computed by subtracting PEP, RSA, resp-RSA, HR, SBP, and DBP baseline scores from the associated task scores. These reactivity scores reflected cardiovascular responses to the speech-in-noise task and constituted our final dependent variables. Given that we also had a baseline measure of self-reported fatigue, we employed the same change-score approach to the fatigue measure. That is, we used the fatigue score of the preceding measure as baseline to quantify the specific fatigue response induced by a certain listening demand level.

### 2.5 Statistical Analysis

We applied a priori planned contrasts (Rosenthal & Rosnow, 1985) to test our hypotheses about the impact of listening demand on PEP and RSA response. We modelled the expected quadratic relationships using contrast weights combing standard quadratic polynomial contrast weights with the prediction of equal response size in the low-demand and impossible-demand conditions. The resulting contrast weights were +5 (low demand), +1 (moderate demand), -11 (high demand), and +5 (impossible demand) for PEP reactivity and +7 (low demand), -5 (moderate demand), -9 (high demand), and +7 (impossible demand) for RSA and resp-RSA reactivity. To examine whether the quadratic relationship hypothesis provided a better explanation of the data as the sawtooth relationship model—linear increase across the three possible demand conditions and disengagement in the impossible condition—predicted by motivational intensity theory, we compared the quadratic model with the sawtooth model (contrast weights: -3 in the low-demand, +1 in the moderate demand, +5 in the high-demand, and -3 in the impossible-demand conditions; e.g., Richter et al., 2008) using Bayes Factors (Masson, 2011; Richter, 2016b). The observed Bayes Factors were interpreted according to Andraszewicz et al. (2014).

We also used planned contrasts to model predictions for HR, SBP, DBP, self-reported effort, fatigue, and performance. Given that HR, SBP, and DBP constitute cardiovascular measures that are influenced by the activity of both branches of the ANS, we used the standard set of contrast weights modelling the sawtooth pattern predicted by motivational intensity theory (Richter et al., 2008). We used the same contrast weights to examine the impact of listening demand on self-reported effort. Self-reported fatigue and task performance—the number of correctly answered multiple-choice questions—were analysed with a standard linear contrast modelling increased fatigue and decreased performance with increasing listening demand. Given that all these predictions were directional and effects in the opposite direction uninterpretable or uninteresting, we employed one-tailed tests (Hales, 2016; Kimmel, 1957). Moreover, to prevent type-I (alpha) error inflation, we only conducted these planned contrasts and refrained from using p-value based tests to explore any effects that we had not predicted.

## 3. Results

### 3.1 Physiological Baselines

Table 1 displays condition means and standard errors of PEP, RSA, resp-RSA, SBP, DBP, and HR baseline scores. Repeated measures correlations (Bakdash & Marusich, 2017) between all assessed cardiovascular measures, performance, and self-report measures can be found in Table 2. Respiration rate (in cycles per minute) was 17.74 (*SE* = 0.49) in the baseline period preceding the low-demand condition, 17.39 (*SE* = 0.44) preceding the moderate-demand condition, 17.64 (*SE* = 0.45) preceding the high-demand condition, and 17.73 (*SE* = 0.45) preceding the impossible condition.

### 3.2 Physiological Reactivity

Table 3 shows condition means and standard errors of all cardiovascular measures. Respiration rate during task performance was as follows: 18.62 (*SE* = 0.48) in the low-demand condition, 18.56 (*SE* = 0.46) in the moderate-demand condition, 18.67 (*SE* = 0.48) in the high-demand condition, and 18.55 (*SE* = 0.42) in the impossible condition.

The planned contrast was significant for PEP, $t(132) = 2.05$, $p = .02$, $r_{\text{contrast}} = .30$, supporting the predicted relationship between listening demand and SNS response. However, the contrast was not significant for RSA, $t(132) = 1.58$, $p = .06$, $r_{\text{contrast}} = .23$, or resp-RSA reactivity, $t(102) = 1.18$, $p = .12$, $r_{\text{contrast}} = .20$, providing no evidence for the predicted effect of listening demand on PNS response. Figures 3 and 4 show the observed patterns of PEP and RSA reactivity. Comparing the predicted quadratic relationship model with the standard sawtooth model did not strongly favour any of the two models: $BF = 0.51$ for PEP, $BF = 1.51$ for RSA, and $BF = 1.25$ for resp-RSA. The planned contrast was not significant for HR, $t(132) = 0.29$, $p = .39$, $r_{\text{contrast}} = .04$, SBP, $t(132) = 0.62$, $p = .27$, $r_{\text{contrast}} = .09$, or DBP, $t(132) = 0.91$, $p = .18$, $r_{\text{contrast}} = .14$.

### 3.3 Task Performance and Self-reports

Table 4 displays condition means and standard errors of all self-reports and task performance. Significant linear contrasts for task performance, $t(132) = 22.60$, $p < .001$, $r_{\text{contrast}} = .96$, and self-reported task demand, $t(132) = 11.76$, $p < .001$, $r_{\text{contrast}} = .87$, suggested a successful manipulation of listening demand. Self-reported fatigue displayed the same linear effect of task demand, $t(132) = 4.03$, $p < .001$, $r_{\text{contrast}} = .52$. Self-reported effort showed the expected increase over the three possible demand levels and the decrease in the impossible demand condition, $t(132) = 6.81$, $p < .001$, $r_{\text{contrast}} = .72$.

## 4. Discussion

The observed PEP reactivity pattern provided support for the predicted impact of listening demand on cardiac SNS activity: Pre-ejection period reactivity increased across the three possible listening demand levels and was low if participants were asked to perform an impossible speech-in-noise task. The absence of parallel decreases in DBP and HR suggests that the observed PEP effects indeed reflected changes in underlying sympathetic activity and not changes in pre-load—which would have been indicated by a parallel decrease in HR (Obrist, 1981)—or after-load—which parallel decreases in DBP would have suggested (Sherwood et al., 1990). However, our findings for RSA and resp-RSA failed to provide evidence for the expected relationship between listening demand and

PNS activity: even if the effect sizes were moderate, the planned contrasts were not significant. However, it may be valuable to note that the effect size for RSA was only minimally different from the effect size observed for PEP. Nevertheless, our data only provided conclusive evidence for the hypothesised relationship between listening demand and cardiac SNS activity, not for the relationship between listening demand and cardiac PNS activity.

Interestingly our results summarise in this regard the existing studies on listening effort and motivational intensity theory that examined the activity of both ANS branches. As discussed in the introduction section, these studies consistently found evidence for demand effects on SNS activity (e.g., Chatelain, Silvestrini, & Gendolla, 2016; e.g., Mackersie & Calderon-Moultrie, 2016; Mackersie & Kearney, 2017; Mazeres, Brinkmann, & Richter, 2019; Richter et al., 2008; Seeman & Sims, 2015), but the evidence for effects on PNS activity has been mixed. Some studies found significant effects (e.g., Mackersie & Calderon-Moultrie, 2016; Seeman & Sims, 2015) whereas others have not (e.g., Mackersie & Kearney, 2017; Silvia et al., 2016). The available literature unfortunately does not answer the question whether this variability of PNS effects is due to a weaker association between task demand and PNS response or due to measure-related issues. In comparison to RSA, PEP has the advantage that there are only two main confounding variables—pre-load and after-load (Sherwood et al., 1990)—that may mask or mimic SNS effects on PEP. RSA is influenced by a broader range of variables, which threaten its sensitivity as an indicator of parasympathetic activity (Berntson et al., 1997; Grossman & Taylor, 2007). For instance, changes in respiration frequency and tidal volume may alter RSA without any underlying change in PNS activity.

Even if the PEP data provided strong support for the impact of listening demand on cardiac SNS activity, it is important to note that the postulated model—predicting a quadratic relationship between listening demand and cardiac SNS activity up to the demand level where individuals disengage—did not perform better than the standard motivational intensity theory model—assuming a linear relationship for the range of possible demand levels. The Bayes Factors comparing the two models did not favour our model for PEP reactivity and did also not provide conclusive evidence in favour of it for RSA or resp-RSA reactivity. An inspection of Figures 2 and 3 reveals that the lack of strong evidence for the predicted quadratic relationship between listening demand and PEP reactivity was due to the reactivity in the moderate demand condition being greater than predicted. Moreover, a lack of sensitivity of our experimental design for detecting differences between the two models may have contributed to the lack of conclusive evidence. In our design, the main difference between the two models was the predicted relative distance between the moderate demand condition and the low and high demand conditions. The linear model predicted that the difference in reactivity between the low and moderate demand conditions equals the difference between the moderate and high demand conditions, whereas the quadratic model predicted a smaller difference in PEP reactivity—or a larger difference in the case of RSA reactivity—between the low and moderate demand conditions than between the moderate and high demand conditions. That is, the relative performance of the two models was determined by the observed reactivity in only one of the four demand conditions: the moderate demand condition. Comparing the models in designs that include more than three possible demand levels will enable a better differentiation between our quadratic model and the standard sawtooth model.

It is important to highlight the crucial role of the task demand calibration procedure. The contrast weights that we used to model the expected quadratic relationships assumed equal intervals between the low, moderate, and high demand conditions. That is, they relied on participants perceiving the difference in demand between the low and moderate condition to be the same as the difference between the moderate and difficult condition. If the verbal labels—low, moderately difficult, and difficult—that we used to identify the SNR levels associated with low, moderate, and high demand were not suitable to create equidistant demand levels, our contrast weights would not have been appropriate. For instance, if the actual difference in perceived demand was larger between the low and moderate demand conditions than between the moderate and high demand conditions, a larger contrast weight difference between the low and moderate demand conditions and a smaller contrast weight difference between the moderate and high demand conditions would have been more appropriate. However, this problem seems to be innate to any calibration of subjective demand levels: The calibration will always depend on the employed verbal labels. Alternative demand calibration strategies that are common in listening effort research like using equal SNR differences (e.g., Ohlenforst et al., 2018; Plain et al., 2020) or intelligibility levels (e.g., Koelewijn, Zekveld, Lunner, & Kramer, 2018a; Wendt, Koelewijn, Ksiazek, Kramer, & Lunner, 2018) do not prevent this problem because they can also not guarantee that differences in perceived demand between consecutive demand levels are equidistant.

In addition to demonstrating the impact of listening demand on cardiac sympathetic response, our data also provided evidence for disengagement under conditions of very high, impossible listening demand. Empirical work on motivational intensity theory has frequently examined whether individuals disengage if task success is impossible or not worth the required effort (see Stanek & Richter, 2016, for a meta-analytic review of 40 studies) but psychophysiological work on listening effort started only recently to acknowledge that the relationship between listening demand and effort may have an upper limit (Ohlenforst et al., 2018; Ohlenforst et al., 2017; Richter, 2016a; Wendt et al., 2018; Winn, Wendt, Koelewijn, & Kuchinsky, 2018; Zekveld & Kramer, 2014; Zhang, Siegle, McNeil, Pratt, & Palmer, 2019). Interestingly all listening effort studies that showed disengagement at extremely high (impossible) demand levels used pupil dilation as indicator of listening effort (Ohlenforst et al., 2018; Ohlenforst et al., 2017; Wendt et al., 2018; Zekveld & Kramer, 2014). Our findings replicate and extend these studies by demonstrating that disengagement in listening tasks is also observable on cardiac sympathetic responses.

The next important step to develop a comprehensive understanding of the psychophysiology of listening effort seems to build on the approach that Seeman and Sims (2015) and Mackersie and colleagues begun (Mackersie & Calderon-Moultrie, 2016; Mackersie & Kearney, 2017) and to always assesses the activity of both ANS branches if peripheral psychophysiological correlates of listening effort are examined. Our study extended their work by focusing on SNS and PNS impact on one and the same organ, and by examining listening demand effects across more than two task demand levels. Given that the pupil is also innervated by both ANS systems, it would be easy to adopt this approach also in listening effort studies that use pupillometry—probably the most frequently assessed psychophysiological correlate of listening effort. Wang et al. (2018) already demonstrated how the method suggested by Steinhauer, Siegle, Condray, and Pless (2004) for the differentiation of SNS and PNS contribution to pupil dilation can be used in listening tasks. Future pupillometric listening effort studies should follow their example and aim to separate SNS and PNS responses. If future listening effort studies assessing peripheral physiological correlates of listening effort consistently examined the individual contribution of both ANS branches, we would probably have in a few years a good understanding of the ANS mechanisms underlying effortful listening.

## 5. Conclusion

The presented findings demonstrated that myocardial sympathetic activity, but not parasympathetic activity, increased as a function of the demand of our speech-in-noise task if task success was possible. They also revealed that both sympathetic and parasympathetic activity were low if it was impossible to understand the speech. Our data thus illustrate that it is important to acknowledge that the relationship between listening demand and effort is more complex than a simple monotonic relationship. If listening demand is too high, individuals may give up and not invest any effort in understanding speech. Moreover, listening effort research should focus on myocardial sympathetic activity when examining physiological correlates of listening effort and might consider sympathetic activity as a potential candidate for an indicator of listening effort.

## Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors, but was supported by a Liverpool John Moores University Faculty of Science PhD studentship.

## Acknowledgements

We are grateful to Ruth Ogden and Ralph Pawling for comments on an early version of this article.

## References

Alhanbali, S., Dawes, P., Lloyd, S., & Munro, K. J. (2017). Self-reported listening-related effort and fatigue in hearing-impaired adults. *Ear and Hearing, 38*(1), e39–e48. doi.org/10.1097/AUD.0000000000000361

Alhanbali, S., Dawes, P., Lloyd, S., & Munro, K. J. (2017). Self-reported listening-related effort and fatigue in hearing-impaired adults. *Ear and Hearing, 38*(1), e39-e48.

Alhanbali, S., Dawes, P., Millman, R. E., & Munro, K. J. (2019). Measures of listening effort are multidimensional. *Ear and Hearing, 40*(5), 1084-1097. doi:10.1097/AUD.0000000000000697

Andraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R., Verhagen, J., & Wagenmakers, E.-J. (2014). An introduction to bayesian hypothesis testing for management research. *Journal of Management, 41*(2), 521-543. doi:10.1177/0149206314560412

Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology, 8*, 456. doi:10.3389/fpsyg.2017.00456

Bernarding, C., Strauss, D. J., Hannemann, R., Seidler, H., & Corona-Strauss, F. I. (2017). Neurodynamic evaluation of hearing aid features using eeg correlates of listening effort. *Cognitive Neurodynamics, 11*(3), 203-215. doi:10.1007/s11571-017-9425-5

Berntson, G. G., Bigger, J. T., Jr., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., . . . van der Molen, M. W. (1997). Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology, 34*(6), 623-648. doi:10.1111/j.1469-8986.1997.tb02140.x.

Berntson, G. G., Cacioppo, J. T., & Quigley, K. S. (1991). Autonomic determinism: The modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. *Psychological Review, 98*(4), 459-487. doi:10.1037//0033-295X.98.4.459

Boyce, W. T., & Ellis, B. J. (2005). Biological sensitivity to context: I. An evolutionary-developmental theory of the origins and functions of stress reactivity. *Development and Psychopathology, 17*(2), 271-301. doi:10.1017/s0954579405050145

Brehm, J. W., & Self, E. A. (1989). The intensity of motivation. *Annual Review of Psychology, 40*(1), 109-131. doi:10.1146/annurev.ps.40.020189.000545

Brickenkamp, K. (2002). *Test d2 aufmerksamkeits-belastungs-test, manual [the d2 test of attention]*: Hogrefe.

Chatelain, M., Silvestrini, N., & Gendolla, G. H. E. (2016). Task difficulty moderates implicit fear and anger effects on effort-related cardiac response. *Biological Psychology, 115*, 94-100. doi:10.1016/j.biopsycho.2016.01.014

Dawson, M. E., Schell, A. M., & Filion, D. L. (2017). The electrodermal system. In *Handbook of psychophysiology* (pp. 217-243).

de Geus, E. J. C., Willemsen, G. H. M., Klaver, C. H. A. M., & van Doornen, L. J. P. (1995). Ambulatory measurement of respiratory sinus arrhythmia and respiration rate. *Biological Psychology, 41*(3), 205-227. doi:10.1016/0301-0511(95)05137-6

Esler, M., Jennings, G., Lambert, G., Meredith, I., Horne, M., & Eisenhofer, G. (1990). Overflow of catecholamine neurotransmitters to the circulation: Source, fate, and functions. *Physiological Reviews, 70*(4), 963-985. doi:10.1152/physrev.1990.70.4.963

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175-191. doi:10.3758/bf03193146

Francis, A. L., & Love, J. (2020). Listening effort: Are we measuring cognition or affect, or both? *Wiley Interdisciplinary Reviews Cognitive Science, 11*(1), e1514. doi:10.1002/wcs.1514

Gendolla, G. H., & Krusken, J. (2001). The joint impact of mood state and task difficulty on cardiovascular and electrodermal reactivity in active coping. *Psychophysiology, 38*(3), 548-556. doi:10.1017/S0048577201000622

Gendolla, G. H. E., Wright, R. A., & Richter, M. (2019). Advancing issues in motivation intensity research: Updated insights from the cardiovascular system. In R. M. Ryan (Ed.), *The oxford handbook of human motivation* (pp. 373-392). New York, NY: Oxford University Press.

Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biol Psychol, 74*(2), 263-285. doi:10.1016/j.biopsycho.2005.11.014

Hales, A. H. (2016). Does the conclusion follow from the evidence? Recommendations for improving research. *Journal of Experimental Social Psychology, 66*, 39-46. doi:10.1016/j.jesp.2015.09.011

Harper, K. L., Eddington, K. M., & Silvia, P. J. (2016). Perfectionism and effort-related cardiac activity: Do perfectionists try harder? *PLoS One, 11*(8), e0160340. doi:10.1371/journal.pone.0160340

Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183): Elsevier.

Hernando, A., Lazaro, J., Gil, E., Arza, A., Garzon, J. M., Lopez-Anton, R., . . . Bailon, R. (2016). Inclusion of respiratory frequency information in heart rate variability analysis for stress assessment. *IEEE Journal of Biomedical and Health Informatics, 20*(4), 1016-1025. doi: 10.1109/JBHI.2016.2553578

Hornsby, B. W. Y. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear and Hearing, 34*(5), 523-534.

Houtveen, J. H., Groot, P. F., & de Geus, E. J. (2006). Validation of the thoracic impedance derived respiratory signal using multilevel analysis. *Int J Psychophysiol, 59*(2), 97-106. doi:10.1016/j.ijpsycho.2005.02.003

Kelsey, R. M., & Guethlein, W. (1990). An evaluation of the ensemble averaged impedance cardiogram. *Psychophysiology, 27*(1), 24-33. doi:10.1111/j.1469-8986.1990.tb02173.x

Kimmel, H. D. (1957). Three criteria for the use of one-tailed tests. *Psychological Bulletin, 54*(4), 351-353. doi:10.1037/h0046737

Koelewijn, T., Zekveld, A. A., Lunner, T., & Kramer, S. E. (2018a). The effect of reward on listening effort as reflected by the pupil dilation response. *Hearing Research, 367*, 106-112. doi:10.1016/j.heares.2018.07.011

Koelewijn, T., Zekveld, A. A., Lunner, T., & Kramer, S. E. (2018b). The effect of reward on listening effort as reflected by the pupil dilation response. *Hear Res, 367*, 106-112. doi:10.1016/j.heares.2018.07.011

Lippman, N., Stein, K. M., & Lerman, B. B. (1994). Comparison of methods for removal of ectopy in measurement of heart rate variability. *American Journal of Physiology: Heart and Circulatory Physiology, 267*(1 Pt 2), H411-418. doi:10.1152/ajpheart.1994.267.1.H411

Litvack, D. A., Oberlander, T. F., Carney, L. H., & Saul, J. P. (1995). Time and frequency domain methods for heart rate variability analysis: A methodological comparison. *Psychophysiology, 32*(5), 492-504. doi: 10.1111/j.1469-8986.1995.tb02101.x

Llabre, M. M., Spitzer, S. B., Saab, P. G., Ironson, G. H., & Schneiderman, N. (1991). The reliability and specificity of delta versus residualized change as measures of cardiovascular reactivity to behavioral challenges. *Psychophysiology, 28*(6), 701-711. doi:10.1111/j.1469-8986.1991.tb01017.x

Mackersie, C. L., & Calderon-Moultrie, N. (2016). Autonomic nervous system reactivity during speech repetition tasks: Heart rate variability and skin conductance. *Ear and Hearing, 37 Suppl 1*, 118S-125S. doi: 10.1097/AUD.0000000000000305

Mackersie, C. L., & Cones, H. (2011). Subjective and psychophysiological indexes of listening effort in a competing-talker task. *Journal of the American Academy of Audiology, 22*(2), 113-122. doi:10.3766/jaaa.22.2.6.

Mackersie, C. L., & Kearney, L. (2017). Autonomic nervous system responses to hearing-related demand and evaluative threat. *American Journal of Audiology, 26*(3S), 373-377. doi:10.1044/2017_AJA-16-0133

Masson, M. E. (2011). A tutorial on a practical bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods, 43*(3), 679-690. doi:10.3758/s13428-010-0049-5

Mazeres, F., Brinkmann, K., & Richter, M. (2019). Implicit achievement motive limits the impact of task difficulty on effort-related cardiovascular response. *Journal of Research in Personality, 82*. doi:10.1016/j.jrp.2019.06.012

McArdle, W. D., Katch, F. I., & Katch, V. L. (2010). *Exercise physiology: Nutrition, energy, and human performance* (7th ed.): Lippincott Williams & Wilkins.

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., & Amitay, S. (2014). Listening effort and fatigue: What exactly are we measuring? A british society of audiology cognition in hearing special interest group 'white paper'. *International Journal of Audiology, 53*(7), 433-440. doi:10.3109/14992027.2014.890296

Michael, S., Graham, K. S., & Davis, G. M. O. (2017). Cardiac autonomic responses during exercise and post-exercise recovery using heart rate variability and systolic time intervals-a review. *Frontiers in Physiology, 8*, 301. doi:10.3389/fphys.2017.00301

Miles, K., McMahon, C., Boisvert, I., Ibrahim, R., de Lissa, P., Graham, P., & Lyxell, B. (2017). Objective assessment of listening effort: Coregistration of pupillometry and eeg. *Trends in Hearing, 21*, 2331216517706396. doi:10.1177/2331216517706396

Nesse, R. M., Bhatnagar, S., & Ellis, B. (2016). Evolutionary origins and functions of the stress response system. In *Stress: Concepts, cognition, emotion, and behavior* (pp. 95-101).

Nesse, R. M., Bhatnagar, S., & Young, E. A. (2007). Evolutionary origins and functions of the stress response. In *Encyclopedia of stress* (2nd ed., pp. 965-970): Elsevier.

Newlin, D. B., & Levenson, R. W. (1979). Pre-ejection period: Measuring beta-adrenergic influences upon the heart. *Psychophysiology, 16*(6), 546-553. doi:10.1111/j.1469-8986.1979.tb01519.x

Obrist, P. A. (1981). *Cardiovascular psychophysiology: A perspective*: Plenum Press.

Obrist, P. A., Light, K. C., James, S. A., & Strogatz, D. S. (1987). Cardiovascular responses to stress: I. Measures of myocardial response and relationship to high resting systolic pressure and parental hypertension. *Psychophysiology, 24*(1), 65-78. doi:10.1111/j.1469-8986.1987.tb01864.x

Ohlenforst, B., Wendt, D., Kramer, S. E., Naylor, G., Zekveld, A. A., & Lunner, T. (2018). Impact of snr, masker type and noise reduction processing on sentence recognition performance and listening effort as indicated by the pupil dilation response. *Hearing Research, 365*, 90-99. doi:10.1016/j.heares.2018.05.003

Ohlenforst, B., Zekveld, A. A., Lunner, T., Wendt, D., Naylor, G., Wang, Y., . . . Kramer, S. E. (2017). Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hearing Research, 351*, 68-79. doi:10.1016/j.heares.2017.05.012

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., . . . Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (fuel). *Ear Hear, 37 Suppl 1*, 5S-27S. doi:10.1097/AUD.0000000000000312

Plain, B., Richter, M., Zekveld, A. A., Lunner, T., Bhuiyan, T., & Kramer, S. E. (2020). Investigating the influences of task demand and reward on cardiac pre-ejection period (pep) reactivity during a speech-in-noise task. *Ear and Hearing*. doi: 10.1097/AUD.0000000000000971

Richter, M. (2010a). Bluebox 2 (Version v1.22). University of Geneva.

Richter, M. (2010b). Pay attention to your manipulation checks! Reward impact on cardiac reactivity is moderated by task context. *Biological Psychology, 84*(2), 279-289. doi:10.1016/j.biopsycho.2010.02.014

Richter, M. (2013). A closer look into the multi-layer structure of motivational intensity theory. *Social and Personality Psychology Compass, 7*(1), 1-12. doi:10.1111/spc3.12007

Richter, M. (2016a). The moderating effect of success importance on the relationship between listening demand and listening effort. *Ear and Hearing, 37*, 111S-117S. doi:10.1097/AUD.0000000000000295

Richter, M. (2016b). Residual tests in the analysis of planned contrasts: Problems and solutions. *Psychological Methods, 21*(1), 112-120. doi:10.1037/met0000044

Richter, M., Friedrich, A., & Gendolla, G. H. E. (2008). Task difficulty effects on cardiac activity. *Psychophysiology, 45*(5), 869-875. doi:10.1111/j.1469-8986.2008.00688.x

Richter, M., Gendolla, G. H. E., & Wright, R. A. (2016). Three decades of research on motivational intensity theory. In *Advances in motivation science* (Vol. 3, pp. 149-186). Cambridge, MA: Academic Press.

Robinson, B. F., Epstein, S. E., Beiser, G. D., & Braunwald, E. (1966). Control of heart rate by the autonomic nervous system. Studies in man on the interrelation between baroreceptor mechanisms and exercise. *Circulation Research, 19*(2), 400-411. doi:10.1161/01.RES.19.2.400

Rosenthal, R., & Rosnow, R. L. (1985). *Contrast analysis. Focused comparisons in the analysis of variance*. Cambridge, England: Cambridge University Press.

Seeman, S., & Sims, R. (2015). Comparison of psychophysiological and dual-task measures of listening effort. *Journal of Speech, Language, and Hearing Research, 58*(6), 1781-1792. doi:10.1044/2015_JSLHR-H-14-0180

Seppa, V. P., Viik, J., & Hyttinen, J. (2010). Assessment of pulmonary flow using impedance pneumography. *IEEE Transactions on Biomedical Engineering, 57*(9), 2277-2285. doi:10.1109/TBME.2010.2051668

Sherwood, A., Allen, M. T., Fahrenberg, J., Kelsey, R. M., Lovallo, W. R., & van Doornen, L. J. (1990). Methodological guidelines for impedance cardiography. *Psychophysiology, 27*(1), 1-23. doi:10.1111/j.1469-8986.1990.tb02171.x

Silvia, P. J., Beaty, R. E., Nusbaum, E. C., Eddington, K. M., & Kwapil, T. R. (2014). Creative motivation: Creative achievement predicts cardiac autonomic markers of effort during divergent thinking. *Biol Psychol, 102*, 30-37. doi:10.1016/j.biopsycho.2014.07.010

Silvia, P. J., Eddington, K. M., Beaty, R. E., Nusbaum, E. C., & Kwapil, T. R. (2013). Gritty people try harder: Grit and effort-related cardiac autonomic activity during an active coping challenge. *International Journal of Psychophysiology, 88*(2), 200-205. doi:10.1016/j.ijpsycho.2013.04.007

Silvia, P. J., Mironovova, Z., McHone, A. N., Sperry, S. H., Harper, K. L., Kwapil, T. R., & Eddington, K. M. (2016). Do depressive symptoms "blunt" effort? An analysis of cardiac engagement and withdrawal for an increasingly difficult task. *Biological Psychology, 118*, 52-60. doi:10.1016/j.biopsycho.2016.04.068

Silvia, P. J., Nusbaum, E. C., Eddington, K. M., Beaty, R. E., & Kwapil, T. R. (2014). Effort deficits and depression: The influence of anhedonic depressive symptoms on cardiac autonomic activity during a mental challenge. *Motivation and Emotion, 38*(6), 779-789. doi:10.1007/s11031-014-9443-0

Skytioti, M., Sovik, S., & Elstad, M. (2017). Respiration-related cerebral blood flow variability increases during control-mode non-invasive ventilation in normovolemia and hypovolemia. *European Journal of Applied Physiology, 117*(11), 2237-2249. doi:10.1007/s00421-017-3711-0

Stanek, J., & Richter, M. (2016). Evidence against the primacy of energy conservation: Exerted force in possible and impossible handgrip tasks. *Motivation Science, 2*(1), 49-65. doi:10.1037/mot0000028

Steinhauer, S. R., Siegle, G. J., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology, 52*(1), 77-86. doi:10.1016/j.ijpsycho.2003.12.005

Strand, J. F., Brown, V. A., Merchant, M. B., Brown, H. E., & Smith, J. (2018). Measuring listening effort: Convergent validity, sensitivity, and links with cognitive and personality measures. *Journal of Speech, Language, and Hearing Research, 61*(6), 1463-1486. doi:10.1044/2018_JSLHR-H-17-0257

Task Force of the European Society of Cardiology and the North American Society of Pacing Electrophysiology. (1996). Heart rate variability. *Circulation, 93*(5), 1043-1065. doi:10.1161/01.CIR.93.5.1043

Venables, L., & Fairclough, S. H. (2009). The influence of performance feedback on goal-setting and mental effort regulation. *Motivation and Emotion, 33*(1), 63-74. doi:10.1007/s11031-008-9116-y

Wang, Y., Kramer, S. E., Wendt, D., Naylor, G., Lunner, T., & Zekveld, A. A. (2018). The pupil dilation response during speech perception in dark and light: The involvement of the parasympathetic nervous system in listening effort. *Trends in Hearing, 22*. doi:10.1177/2331216518816603

Wendt, D., Koelewijn, T., Ksiazek, P., Kramer, S. E., & Lunner, T. (2018). Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test. *Hear Res, 369*, 67-78. doi:10.1016/j.heares.2018.05.006

White, D. W., & Raven, P. B. (2014). Autonomic neural control of heart rate during dynamic exercise: Revisited. *Journal of Physiology, 592*(12), 2491-2500. doi:10.1113/jphysiol.2014.271858

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: The processing of degraded speech depends critically on attention. *The Journal of Neuroscience, 32*(40), 14010-14021. doi: 10.1523/JNEUROSCI.1528-12.2012

Williams, M. (Writer). (2012). Kingdom of plants 3d. In A. Productions (Producer).

Winn, M. B., Wendt, D., Koelewijn, T., & Kuchinsky, S. E. (2018). Best practices and advice for using pupillometry to measure listening effort: An introduction for those who want to get started. *Trends in Hearing, 22*, 2331216518800869. doi:10.1177/2331216518800869

Wright, R. A. (1996). Brehm's theory of motivation as a model of effort and cardiovascular response. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 424-453). New York, NY: Guilford Press.

Wright, R. A. (1998). Ability perception and cardiovascular response to behavioral challenge. In M. Kofta, G. Weary, & G. Sedek (Eds.), *Personal control in action: Cognitive and motivational mechanisms* (pp. 197-232). Boston, MA: Springer US.

Wright, R. A. (2008). Refining the prediction of effort: Brehm's distinction between potential motivation and motivation intensity. *Social and Personality Psychology Compass, 2*(2), 682-701. doi:10.1111/j.1751-9004.2008.00093.x

Wright, R. A., & Franklin, J. (2004). Ability perception determinants of effort-related cardiovascular responses: Mood, optimism, and performance resources. In R. A. Wright & J. Greenberg (Eds.), *Motivational analyses of social behavior: Building on jack brehm's contributions to psychology* (pp. 187-204): Lawrence Erlbaum Associates Publishers.

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology, 51*(3), 277-284. doi:10.1111/psyp.12151

Zhang, M., Siegle, G. J., McNeil, M. R., Pratt, S. R., & Palmer, C. (2019). The role of reward and task demand in value-based strategic allocation of auditory comprehension effort. *Hearing Research, 381*, 107775. doi:10.1016/j.heares.2019.107775

**Table 1**

*Means and Standard Errors of Cardiovascular Baselines Scores*

| Variable | Low | | Moderate | | High | | Impossible | |
|---|---|---|---|---|---|---|---|---|
| | *M* | *SE* | *M* | *SE* | *M* | *SE* | *M* | *SE* |
| PEP | 100.92 | 2.12 | 101.42 | 1.99 | 102.41 | 1.96 | 101.78 | 2.01 |
| RSA | 34.83 | 2.63 | 36.89 | 2.81 | 35.66 | 2.83 | 33.93 | 2.57 |
| resp-RSA | 22.77 | 2.37 | 25.26 | 2.67 | 23.89 | 2.50 | 22.53 | 2.38 |
| SBP | 107.89 | 1.33 | 107.75 | 1.51 | 107.97 | 1.36 | 108.27 | 1.43 |
| DBP | 69.07 | 1.06 | 69.26 | 0.98 | 69.71 | 1.04 | 69.60 | 1.15 |
| HR | 72.06 | 1.76 | 71.95 | 1.72 | 71.60 | 1.54 | 71.59 | 1.60 |

*Note.* $n$ = 35 for resp-RSA. $N$ = 45 for all other measures. PEP is in ms, RSA and resp-RSA are in nu, SBP and DBP are in mmHg, and HR is in bpm.

**Table 2**

*Bivariate Correlation Coefficients for Cardiovascular Measures, Performance, and Self-report Measures*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. PEP baseline | — | | | | | | | | | | | | | | |
| 2. RSA baseline | .11 | — | | | | | | | | | | | | | |
| 3. resp-RSA baseline | -.01 | .83 | — | | | | | | | | | | | | |
| 4. SBP baseline | -.21 | -.08 | -.01 | — | | | | | | | | | | | |
| 5. DBP baseline | .03 | -.01 | .07 | .41 | — | | | | | | | | | | |
| 6. HR baseline | .00 | -.29 | -.19 | .23 | .16 | — | | | | | | | | | |
| 7. PEP reactivity | -.31 | -.19 | -.17 | ,08 | .00 | -.12 | — | | | | | | | | |
| 8. RSA reactivity | .04 | -.65 | -.51 | -.04 | -.05 | .18 | .17 | — | | | | | | | |
| 9. resp-RSA reactivity | .15 | -.53 | -.68 | -.08 | -.14 | .16 | .13 | .76 | — | | | | | | |
| 10. SBP reactivity | -.02 | -.03 | -.18 | -.41 | -.08 | .01 | -.03 | .03 | .17 | — | | | | | |
| 11. DBP reactivity | .01 | .02 | -.09 | .04 | -.55 | .04 | .03 | .02 | .10 | .21 | — | | | | |
| 12. HR reactivity | -.06 | .11 | -.02 | .17 | .08 | -.28 | .13 | -.15 | .03 | .17 | .10 | — | | | |
| 13. Performance | .01 | .07 | .05 | .02 | -.05 | .08 | -.09 | .00 | .05 | .13 | .05 | -.02 | — | | |
| 14. Demand | .06 | -.07 | -.05 | .02 | .05 | -.04 | -.10 | .04 | -.02 | .01 | .02 | -.03 | -.50 | — | |
| 15. Effort | .09 | -.04 | -.05 | .01 | .05 | -.02 | -.09 | .02 | -.03 | .05 | -.06 | -.01 | -.30 | .83 | — |
| 16. Fatigue | -.09 | -.04 | -.05 | .09 | -.05 | .00 | .07 | .04 | -.01 | -.05 | .12 | .05 | -.34 | .24 | .20 |

*Note.* *n* = 35 for all correlations involving resp-RSA. *N* = 45 for all other measures. Correlations are repeated measures correlation (rmcorr) coefficients (Bakdash & Marusich, 2017)

**Table 3**

*Means and Standard Errors of Cardiovascular Reactivity Scores*

| Variable | Low | | Moderate | | High | | Impossible | |
|---|---|---|---|---|---|---|---|---|
| | *M* | *SE* | *M* | *SE* | *M* | *SE* | *M* | *SE* |
| PEP | -0.21 | 0.44 | -1.02 | 0.49 | -1.29 | 0.47 | -0.15 | 0.38 |
| RSA | -0.43 | 1.74 | -4.10 | 1.87 | -3.30 | 1.71 | -1.44 | 1.63 |
| resp-RSA | -1.87 | 1.76 | -5.12 | 1.56 | -4.43 | 1.84 | -3.28 | 1.51 |
| SBP | 2.00 | 0.61 | 1.01 | 0.64 | 1.76 | 0.69 | 0.45 | 0.80 |
| DBP | 1.51 | 0.72 | 1.54 | 0.58 | 0.86 | 0.54 | 1.55 | 0.60 |
| HR | 3.04 | 0.54 | 3.72 | 0.55 | 3.40 | 0.51 | 3.58 | 0.68 |

*Note.* $n$ = 35 for resp-RSA. $N$ = 45 for all other measures. PEP is in ms, RSA and resp-RSA are in nu, SBP and DBP are in mmHg, and HR is in bpm.

**Table 4**

*Means and Standard Errors of Task Performance and Self-reported Demand, Effort, and Fatigue*

| Variable | Low | | Moderate | | High | | Impossible | |
|---|---|---|---|---|---|---|---|---|
| | *M* | *SE* | *M* | *SE* | *M* | *SE* | *M* | *SE* |
| Performance | 9.67 | 0.10 | 7.87 | 0.18 | 7.62 | 0.24 | 2.73 | 0.28 |
| Demand | 2.02 | 0.14 | 3.29 | 0.15 | 4.11 | 0.11 | 4.20 | 0.20 |
| Effort | 2.16 | 0.15 | 3.42 | 0.14 | 4.22 | 0.11 | 3.73 | 0.24 |
| Fatigue | -9.11 | 4.78 | 0.35 | 4.22 | 7.60 | 4.75 | 20.47 | 5.96 |

*Note. n* = 45.

## Figure Captions

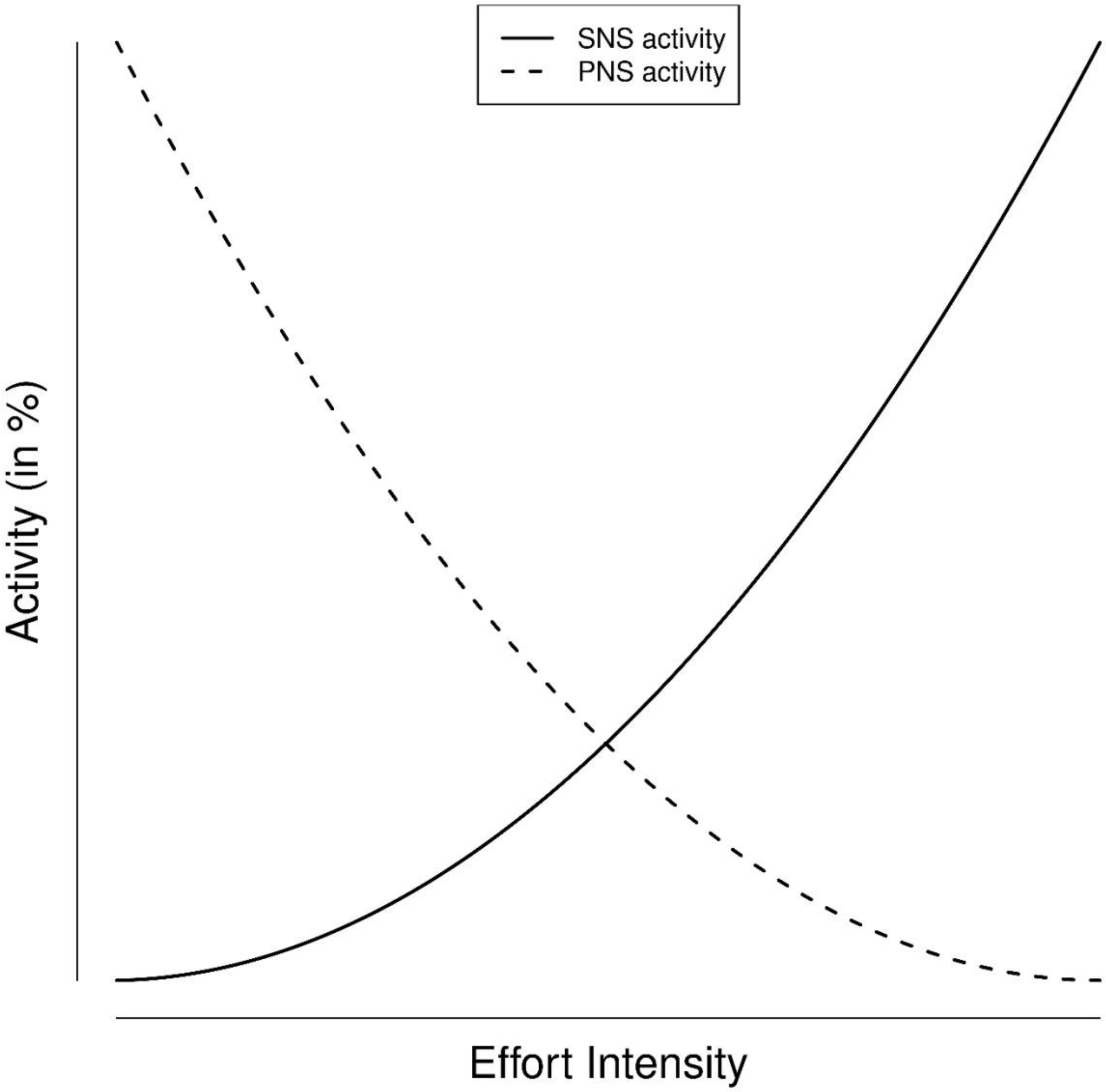
*Figure 1.* Hypothesized relationship between effort intensity and myocardial sympathetic and parasympathetic activity.

*Figure 2*. Predicted PEP and RSA reactivity as a function of listening demand.

*Figure 3*. PEP reactivity—the change from baseline to task—across the four listening demand levels. More negative values reflect an increase in reactivity, and thus increased sympathetic activation.

*Figure 4*. RSA reactivity—the change from baseline to task—across the four listening demand levels. More negative values reflect an increase in reactivity, and thus decreased parasympathetic activation.

**Figure 1**



*Note*. SNS = sympathetic nervous system. PNS = parasympathetic nervous system.

**Figure 2**

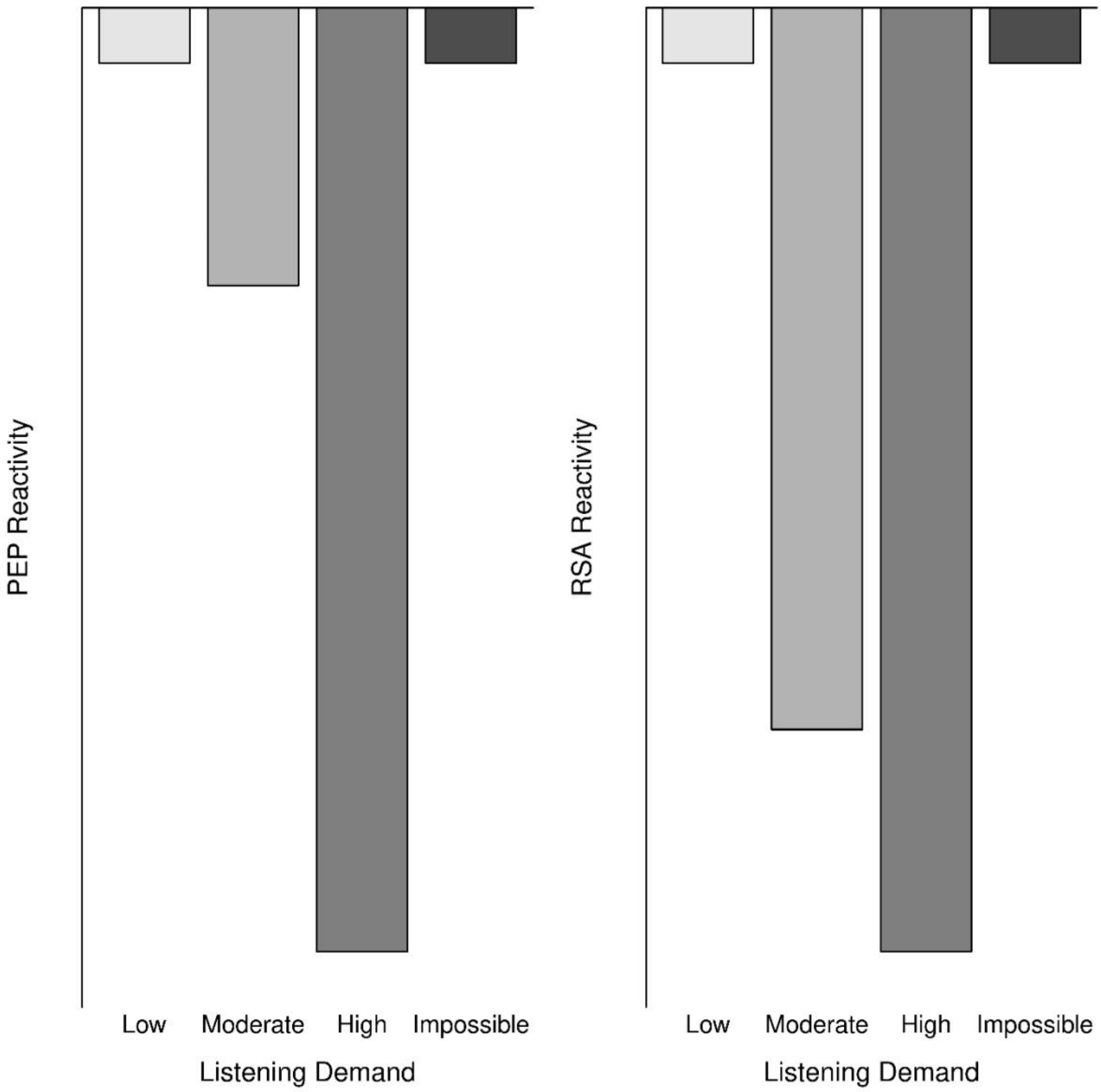
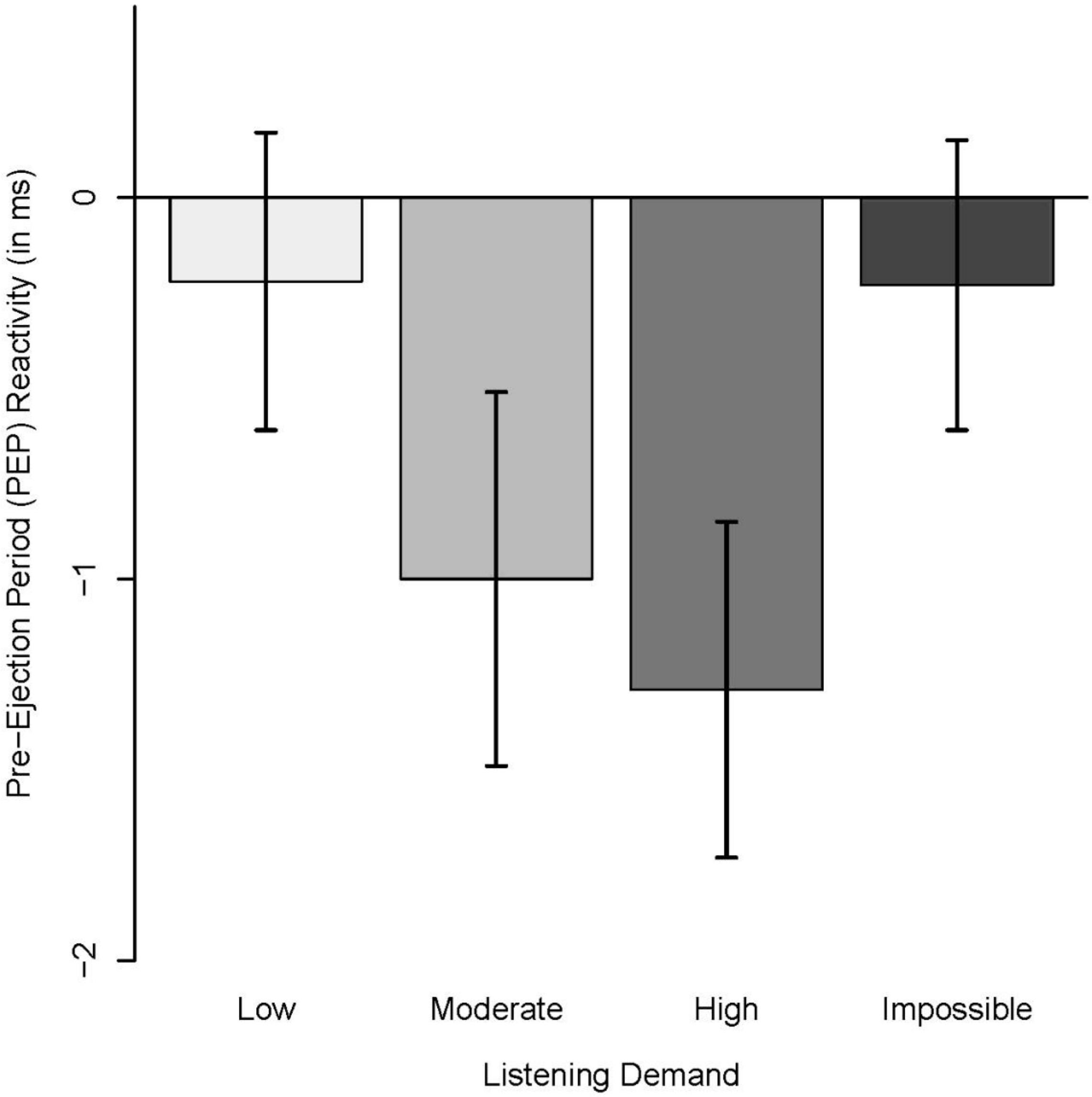
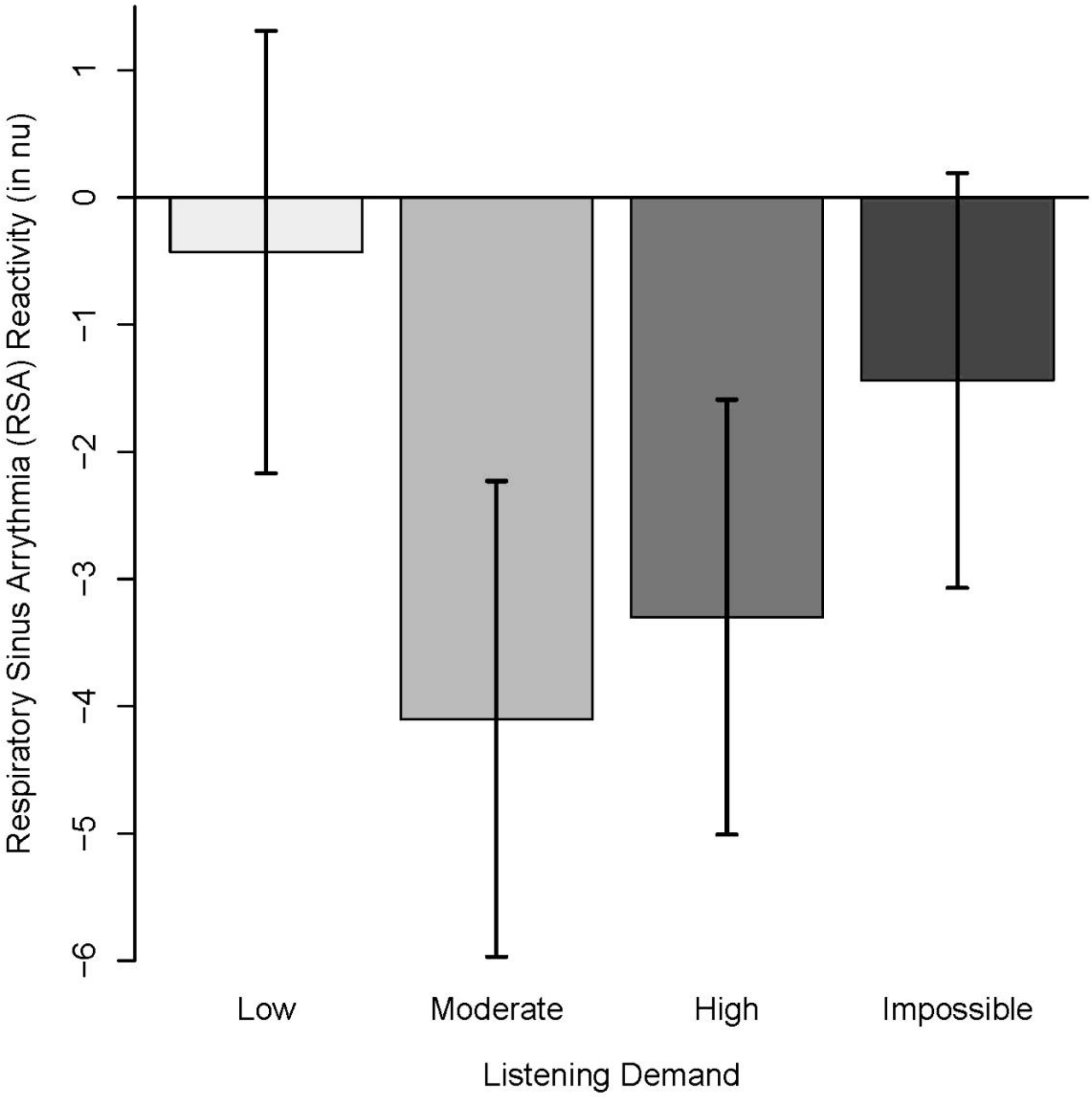

PEP Reactivity
Low
Moderate
High
Impossible
Listening Demand
RSA Reactivity
Low
Moderate
High
Impossible
Listening Demand

**Figure 3**



*Note*. Error bars indicate SEs.

**Figure 4**



*Note*. Error bars indicate SEs.

# Appendix A

The following assumptions underlie the predicted quadratic relationships between (listening) effort and cardiac SNS and PNS activity:

1) Total cardiac ANS activity—the total task-related ANS response caused by increased SNS activity and decreased PNS activity—increases in a linear manner with increases in effort.
2) SNS contribution to total cardiac ANS activity increases in a linear manner with increases in effort, and PNS contribution decreases in a linear manner with increases in effort.
3) At the lowest effort level, SNS activity is close to zero and PNS activity is close to its resting activity.
4) At the highest effort level SNS and PNS contribute each 50% to total cardiac ANS activity. PNS withdrawal is complete at this level and SNS activity is close to its maximum.
5) A unit change in SNS and a unit change in PNS have the same effects on total cardiac ANS response.

## Appendix B

Three examples of the 32-second short stories presented to participants during the speech-in-noise task, as well as the associated comprehension question and 3-option multiple choice responses. The complete set of audio files and lists of all stories can be accessed through the online supplementary materials (https://doi.org/10.24377/LJMU.d.00000087).

| | |
|---|---|
| Short story | Liverpool women's netball club go on a social outing every week, after practicing at the sports centre. This week, the women walked to the station on Friday. They bought three cups of fresh coffee and talked about improving their team strategy for the next game. They considered holding try outs for new team members to improve their capability. |
| Comprehension question | Where did the women go? |
| Multiple choice options | Station / Café / Canteen |
| Short story | Rob works at a garage during the week. He likes his job a lot, but he wishes he had a more physically active role. To try and keep fit, he cycles to work every day. He enjoys it because he rides down the scenic canal path. On Wednesday, Rob decided to sign up for a 5 mile triathlon to encourage himself to cycle more, and to spend more time outdoors. |
| Comprehension question | Where does Rob work? |
| Multiple choice options | Garage / Garden Centre / Golf Course |
| Short story | Students at Wellington School have decided to open a snack stand. The students need fruit to sell at the stand. During lunch time on Monday, one of |

| | |
|---|---|
| | the teachers walked to the supermarket to try to help the students. She picked up three fresh lemons for the snack stand. Then she decided to look for some books to keep in her classroom. |
| Comprehension question | What did the teacher look for? |
| Multiple choice options | Books / Blue-tac / Benches |