The phantoms of the opera—
Stress offstage and stress onstage

Anja-Xiaoxing Cui¹, Negin Motamed Yeganeh¹,
Olga Sviatchenko¹, Thea Leavitt¹,
Taylor McKee¹, Christina Guthier²,
Nancy Hermiston¹ and Lara Boyd¹

Abstract
During opera performance singers deliver vocally demanding roles, follow a conductor, portray emotions of a musical work, act, dance, and engage with costumes, sets and props before an audience. Hence, opera performance is a stressful experience. This study examined different types of stress experiences by measuring the trajectories of 10 opera trainees’ heart rate variability (HRV) during two performances, covering onstage and offstage periods. We explored connections between HRV, self-reported stress measures, and expert-rated difficulty of the performed roles. We discovered that opera trainees had lower HRV and thus experienced greater physiological stress, while onstage compared to offstage periods. In contrast, when asked about performance specific stress, opera trainees self-reported that they felt more nervous when they were offstage. This disconnect between physiological measurement and psychological self-assessment suggests that there are two relevant types of stress for opera performance: psychological stress, which is felt more keenly offstage, and physiological stress, which is greater onstage. Patterns of association between HRV and self-reported measures suggest that HRV is linked to general (not performance-specific) stress. Patterns between self-reported measures suggest that music performance anxiety relates to trait anxiety. Our results indicate specific targets for possible interventions for stress management in opera singers.

Keywords
performance, music performance anxiety, psychophysiology, stress, live performance, heart rate variability

Many invisible processes spur on opera singers to deliver those high notes apart from what the audience sees onstage. From the singers’ perspective, multiple activities have to be coordinated simultaneously: performance of technically and emotionally demanding material, communication with the conductor in the orchestra pit and colleagues on stage, and responding to the

¹University of British Columbia, Vancouver, BC, Canada
²Johannes Gutenberg-University, Mainz, Germany

Corresponding author:
Anja-Xiaoxing Cui, School of Music, University of British Columbia, 6 361 Memorial Rd, Vancouver, BC V6T 1Z2, Canada.
Email: anja.cui@ubc.ca
energy of the audience. Singers may experience “nervousness” or “stage fright” due to these demands.

At the same time, offstage periods present the singers with an opportunity to cognitively evaluate their situation apart from changing costumes or waiting for cues to return onstage. These different demands may lead to distinct forms of stress experiences offstage versus onstage. The research questions this article seeks to address are the following: How do performers’ stress responses vary between different phases of a performance? Are there different kinds of stress experiences? And how do they relate to each other? An opera performance presents a unique opportunity to examine these questions.

It should be noted that opera performers differ from instrumentalists in several ways. For example, while opera performers move around onstage, instrumentalists are often in more static positions when they perform. Furthermore, while opera performers have to memorize music, orchestral musicians usually perform from a score. Thus, performance demands are different for opera performers and instrumentalists, though some issues remain the same.

Anxiety is a common issue in musical performance. Anxiety is understood to have both a state and a trait component (Ladd & Gabrieli, 2015). State anxiety reflects the psychological transient reactions consisting of feelings of apprehension, nervousness, and physiological phenomena such as increased heart rate or respiration (Kennedy et al., 2001). Trait anxiety refers to the stable trend to attend to, experience, and report negative emotions such as fears, worries, and anxiety across many situations (Gidron, 2013). Trait anxiety is frequently linked to anxiety among instrumentalists (Humara, 1999). It may cause musicians to perform more poorly in situations with an evaluative component (Kenny, 2011). Normative data show that 15% of the population is expected to fall in the highly anxious range. In a sample of 48 elite operatic chorus artists, half fell into this range, suggesting a higher prevalence of anxiety among chorus opera artists (Kenny et al., 2004).

Music performance anxiety is characterized by fears of musical presentations, for example, in front of an audience or jury. This complex phenomenon is caused by the interaction of many factors, including cognitions, behaviors, environmental stimuli, and the individual’s experience (Kenny, 2011). Music performance anxiety is highly prevalent among instrumentalists and singers. A recent literature review showed that more than 16% of musicians had indicators of music performance anxiety, with higher prevalence for female musicians (Burin & Osorio, 2017). Higher levels of music performance anxiety are accompanied by physiological, mental, and behavioral symptoms in the performer (Lehmann et al., 2007). Physiological symptoms include changes in heart rate and heart rate variability (HRV).

HRV is a physiological measure of the variation of the time intervals between each heartbeat, called RR intervals. HRV is controlled by the autonomic nervous system, which regulates vital functions in healthy individuals (Kim et al., 2018). However, persistent instigators such as stress can disrupt the autonomic nervous system and thus shift HRV (Pereira et al., 2017). During states of stress, the sympathetic nervous system (a division of the autonomic nervous system) is activated and subsequently suppresses variation between heartbeats, causing a decrease in HRV (Dimsdale, 2008; Kim et al., 2018; Shaffer et al., 2014). Decreased HRV reduces the ability of the body to cope with stress due to impaired regulatory and homeostatic autonomic nervous system functions (Kim et al., 2018). HRV metrics show consistent differences during states of stress (low HRV) compared to relaxed states (high HRV) and thus are a reliable assessment of stress levels in the short-term (Pereira et al., 2017).

Perceived stress is an individual’s perception of the amount of stress at a given period of time. A commonly used clinical tool is the Perceived Stress Scale (Lee, 2012). Self-reported stress measured with this tool positively correlates with serum cortisol levels (Walvekar et al.,
Three types of HRV measures are frequently computed: time-based, frequency-based, and non-linear measures. Time domain parameters include the average interval between heartbeats (Kim et al., 2018; Singh et al., 2018). HRV time domain variables show an inverse relationship between HRV and stress (Wang et al., 1998), such that a significant depression in time domain parameters is associated with higher stress levels (Castaldo et al., 2015; Pereira et al., 2017).

Frequency domain analysis of HRV measures the amount of relative or absolute signal energy detected in various frequency bands (Shaffer & Ginsberg, 2017). Low-frequency and high-frequency energy estimates the activity of the sympathetic and parasympathetic nervous systems, respectively (Kim et al., 2018; Shaffer & Ginsberg, 2017). Increased stress is associated with an imbalance in the autonomic nervous system due to an increase in sympathetic nervous system and decrease in parasympathetic nervous system activity (Kim et al., 2018). Thus, during increased stress, a decrease in the high-frequency band and increase in low-frequency band are expected (Delaney & Brodie, 2000). However, this was not the case in Chanwimalueang et al. (2017), who found lower power in the low-frequency band during the more stressful period.

Non-linear domain analysis of HRV yields additional biomarkers of stress that allow for assessment of the complexity loss hypothesis (Lipsitz & Goldberger, 1992). The complexity loss hypothesis states that humans with healthy physiological functions are able to adapt to the unpredictable changes of human life due to a complex interaction of multiple control mechanisms (Lipsitz & Goldberger, 1992). Individuals under constraints such as illness or stress display lower structural complexity in their physiology, for example, their HRV, than those without. Analysis of the non-linear domain is, therefore, aimed to depict potential changes of this physiological variability (Pincus & Goldberger, 1994).

Structural complexity can be assessed through calculating entropy. Sample entropy is used to indicate regularity in a time series, with increasing values of sample entropy corresponding to higher degrees of irregularity and thus complexity (Richman & Moorman, 2000). These estimates of complexity can provide improved differentiation between healthy and pathological groups in comparison to traditional algorithms (Costa et al., 2002). A recent improvement upon sample entropy, fuzzy entropy can more proficiently measure the regularity of time series (W. Chen et al., 2007). Multiscale sample entropy and multiscale fuzzy entropy examine uncertainty of HRV over multiple time-scales (Costa et al., 2002; Xiong et al., 2010).

Previous research on musicians identified an influence of performance condition on HRV. Differences in low-frequency power were reported during a simulated concert for singers and flutists who were nervous compared to those that were not nervous before the performance (Harmat & Theorell, 2010), and for violinists and flutists between the period preceding a simulated performance and the performance itself (Chanwimalueang et al., 2017). Studies have also found effects of performance conditions on time domain and non-linear domain variables of HRV (Chanwimalueang et al., 2017; Williamon et al., 2013).

Physiological and self-reported measures of stress do not always correspond. For example, in Williamon et al. (2013), the ratio of low-frequency and high-frequency power indicated one performance condition becoming less stressful than another, contradicting the self-report of the pianist who was the case study participant. Similarly, breathing exercises can help instrumentalists and singers regulate frequency domain HRV measures during a period of performance anticipation without decreasing self-reported anxiety (Wells et al., 2012).

A common feature of research into performance stress is the artificial nature of the conditions in which stress is tested, for example, a small audience in a laboratory as a stand-in for a concert-hall experience. Although self-reported anxiety measures can indicate that the manipulation of
anxiety is successful, the degree of similarity to a real-life performance is unknown. Here, we sought to fill this gap in the literature by measuring HRV in opera trainees during two instances of a student opera performance. Unlike most instrumental performances, opera performances often take place in runs, that is, there are repeat performances of the same musical piece.

While this setup prohibits the fine-grained manipulation of performance conditions, we believe the ecological validity of the data is important. By contrasting HRV variables during different naturally presenting performance conditions, for example, onstage versus offstage, we add to the knowledge of performance stress. We further collected information on the singers to explore connections between a variety of self-report measures (e.g., self-reported anxiety, personality, demographics), physiological stress measured through HRV, and expert-rated difficulty of the singers’ roles.

One particular consideration of this study is the impact of controlled breathing on HRV. Changes in breathing patterns are known to affect HRV (Bernardi et al., 2000; Saboul et al., 2013; Weippert et al., 2015). This is especially the case for frequency domain indices (García-González et al., 2000), which further perform worse than temporal and non-linear indices in longitudinal designs (Saboul et al., 2013). Singers control their breathing during singing, thus, active singing periods cannot be directly compared to non-singing periods. We, therefore, studied the effects of performance repetition on those active singing periods using temporal and non-linear indices only, and different types of non-singing periods within performance, that is, offstage periods to onstage silent periods, using all three types of HRV measures.

In summary, the aim of this study was to examine stress responses in opera trainees in an ecologically valid setting. We predicted that given the differential in demands during periods on- versus offstage, and therefore, the differences in potential stress, there would be differences in HRV for opera trainees, while they are silent onstage compared to when they are offstage. Specifically, we hypothesize that physiological stress is higher onstage, whereas psychological stress is higher offstage. We conducted a repeated-measures study, in which stress is measured during different phases of a performance and different performances in the same performer. We further explored the influence of performance repetition and demographic variables on HRV, and assessed whether HRV and self-report measures of stress align in opera singers.

**Methods**

*The opera*

*The Passenger* is a 1968 opera by Mieczysław Weinberg (libretto by Alexander Medvede, 1967). This opera dramatizes the story of a German couple, Walter and Lisa, who are on a transatlantic voyage to South America. Lisa, a former Auschwitz concentration camp guard, believes she recognizes a fellow passenger as Marta, one of the Polish prisoners who had been under her charge at the camp. The opera unfolds on two levels, with some scenes taking place on the ocean liner deck where a horrified Walter listens to his wife’s narration, and the couple interacting with passengers and the crew of the ship. In the other scenes, life at the concentration camp unfolds in all its nightmarish brutality and shows the grueling moments between prisoners and SS soldiers, especially the interactions between Lisa and Marta.

The Passenger consists of two acts. Act 1 contains three scenes and Act 2 contains five scenes and an epilogue. The main roles are Lisa (the former Auschwitz Aufseherin), Marta (the Polish Auschwitz prisoner), Walter (Lisa’s husband), Tadeusz (Marta’s fiancé), and Katja (a Russian Auschwitz prisoner who smuggles a letter). Smaller named roles included Bronka (a Russian Auschwitz prisoner who prays to God), Yvette (a French Auschwitz prisoner who teaches Bronka French), Alte (a long time Auschwitz prisoner who has seen previous prisoners
disappear), Krystina (a Polish Auschwitz prisoner), Vlasta (a Czech Auschwitz prisoner), and Hannah (a Greek Auschwitz prisoner).

Participants
We recruited 10 opera trainees (eight women and two men, age \( M = 26.6 \) years, \( SD = 4.3 \) years) with an experience range of 3–30 past opera productions from the named roles above. Seven of the trainees were graduate students, while the remaining three participants were undergraduate students. This production of the opera was double cast. Singers were not paid to perform in the production, though their performances are graded as part of their curriculum. None of the participants indicated use of medications. This study was reviewed and approved by the Behavioural Research Ethics Board of the University of British Columbia.

Data collection
Data collection took place over the four performances of this production across four consecutive days. Cast A had performances on days 1 and 3 and Cast B had performances on days 2 and 4. HRV was collected using Polar H10 chest straps which were attached and checked by the researchers. The Polar H10 heart rate monitor is validated (Gilgen-Ammann et al., 2019) and used as a criterion measure in other research (Müller et al., 2019; Olstad & Zinner, 2020). The chest straps were paired with iPod touch devices outfitted with the Elite HRV app to allow extraction of RR intervals. The iPods were placed in waist packs hidden underneath costumes. The chest straps and waist packs were applied roughly 1 hour prior to the start of the opera and removed after curtain call.

All participants were invited to fill out online questionnaires regarding their experience and strategies for dealing with performance stress 1 month after their last stage appearance. Responses to the questionnaires were collected through Qualtrics. Participants were asked to indicate whether they felt a difference in nervousness between onstage or offstage periods, and if so where they felt more nervousness. Additionally, participants were asked to rate on a ten-point scale (1 = not nervous at all, 10 = extremely nervous) how nervous they felt before their first, during their first, before their second, and during their second performance.

As part of this questionnaire, we also collected measures of self-esteem (Rosenberg Self-Esteem Scale; Rosenberg, 1965), self-efficacy (New General Self-Efficacy Scale; G. Chen et al., 2001), trait anxiety (State-Trait Anxiety Inventory Y-2; Spielberger, 1983), music performance anxiety (Kenny Music Performance Anxiety Inventory; Kenny, 2011), perceived stress (Perceived Stress Scale; Cohen et al., 1983), and personality (NEO-Five Factor Inventory; Costa & McCrae, 1992). Brief descriptions of the scales are included in the online Supplemental material.

We also collected data from production staff of The Passenger. Seven experts, including singing coaches and members of the production team, filled out an online questionnaire administered through Qualtrics regarding their opinion about each role of the opera to obtain an external evaluation of the varying levels of stress of the different roles.

Data processing
HRV data were exported through the Elite HRV app, which uses a proprietary algorithm to remove artifacts from the RR intervals (Elite HRV, 2020). Video recordings of the performances were analyzed visually by two researchers (AC and NM) to create time tags for the start and end of offstage, onstage, silent onstage, and singing periods. Time tags were created using Adobe
PremierePro and subsequently used to section RR interval data into offstage, onstage, silent onstage, and singing periods, where onstage periods encompass silent and singing periods, and offstage and silent onstage periods are considered non-singing periods.

Three types of HRV measures were computed. Time- and frequency-based HRV measures were computed using code modified from the open source toolbox HRVTool (Vollmer, 2019). Non-linear measures were calculated using code provided by Chanwimalueang et al. (2017). Given the results presented by Chanwimalueang et al. (2017), we chose to focus on the average interval between heartbeats, the low-frequency band, as well as multiscale sample and multi-scale fuzzy entropy as representative measures of the three domains of HRV analysis which may track stress in performers.

Participant questionnaires were processed according to the provided scoring templates with standardized scores for the State-Trait Anxiety Inventory and T scores for the NEO-Five Factor Inventory. Expert questionnaires were used to create a rating of the external stressor by averaging experts’ ratings for each inhabited character according to how “significant,” “stressful,” “emotionally demanding,” and “musically demanding” the roles were, respectively. The four ratings were highly correlated with each other, range of \( r_s(6) = .76–.92, p_s < .05 \). Thus, an average was formed from all four of these ratings to create an overall estimate of how stressful a role is considered.

HRV measures during non-singing periods were compared between performance conditions (offstage vs onstage silent) and performances (Performance 1 vs Performance 2). HRV measures during singing periods were compared between performances (Performance 1 vs Performance 2). To explore the influence of demographic variables, we correlated the latter with HRV. To assess whether HRV and self-reported measures of stress align, we probed their self-reported measures with additional tests.

**Results**

Tables with de-identified, individual data can be found in the online Supplemental material. Because of the small sample size, we report descriptive and non-parametric statistics.

**HRV**

Average HRV for offstage and silent performance conditions is depicted in Table 1, with overall patterns quite consistent across average interval between heartbeats and multiscale sample entropy and multiscale fuzzy entropy, while slightly different for low-frequency. Visual inspection and descriptive statistics suggest effects of performance condition and performance. Indeed, all HRV parameters except for multiscale fuzzy entropy during performance 2 were significantly lower during silent onstage periods compared to the respective offstage periods, Wilcoxon Signed-Rank \( p_s < .05 \), see Table 1 for individual \( p \)-value. However, average interval between heartbeats, multiscale sample entropy, and multiscale fuzzy entropy were not significantly lower during Performance 1 compared to Performance 2. Wilcoxon Signed-Rank \( p_s > .05 \). Means and standard errors of the mean are listed in Table 1.

To assess HRV during active periods of singing, we compared HRV between Performance 1 and Performance 2 using Wilcoxon Signed-Rank tests. There were no significant effects, range of \( Z_s = -1.52 \) to \(-0.17, p_s > .05 \).

**Questionnaire data**

Eight out of nine participants who filled out the questionnaires indicated that there was a difference in how nervous they felt between when they were onstage and when they were offstage.
Seven participants indicated they were more nervous when they were offstage and one participant indicated that being onstage was more stressful. Seven participants indicated that there was a difference in their heart rate between when they were onstage and when they were offstage. Among these participants, five indicated their heart beat faster when they were onstage and two reported that their heart beat faster when they were offstage.

To evaluate the effect of performance condition on the participant’s self-perceived nervousness, we compared their self-report measures between performance phases (before vs during) and performances (Performance 1 vs Performance 2). Self-perceived nervousness was significantly higher before performances compared to during performances, and significantly higher for Performance 1 compared to Performance 2. Means and standard errors of the mean as well as p-values of Wilcoxon Signed-Rank tests are given in Table 2.

The Spearman’s rank-order correlation coefficients between the total music performance anxiety score and various indicators of experience (numbers of opera productions, named roles, big roles, medium roles, small roles, years of opera training, and years of singing training) were consistently negative, range of $r_s(7) = -.79$ to $-.42$, though only the coefficients for numbers of opera productions, named roles, and medium roles were significant, $p_s < .05$. Similarly, correlation coefficients between the total music performance anxiety score and our self-esteem and self-efficacy measures were negative, range of $r_s(7) = -.45$ to $-.09$, though not significant, $p_s > .05$.

Music performance anxiety and negative emotionality were significantly correlated, $r_s(7) = -.80$, $p = .010$, such that low music performance anxiety was associated with higher negative emotionality. Music performance anxiety was not significantly correlated with extraversion, $r_s(7) = -.08$, $p > .05$. Music performance anxiety was also significantly correlated with trait anxiety, $r_s(7) = .94$, $p = .001$, such that high music performance anxiety was associated with high trait anxiety.

### HRV and questionnaire data

We explored the association of a variety of self-report measures and HRV by calculating correlations between the level of comfort of singing in a foreign language, music performance anxiety, perceived stress, and four HRV measures (average interval between heartbeats, low-frequency, multiscale sample entropy, multiscale fuzzy entropy) from six phases (3 performance conditions—offstage, silent, and active—× 2 performances—Performance 1 and Performance 2). There were no significant correlations between measures of HRV and music.
performance anxiety, $p < .05$, except for significant correlations for multiscale sample entropy, $r_s(7) = .73, p = .026$, and multiscale fuzzy entropy, $r_s(7) = .73, p = .026$, during silent conditions in Performance 1.

Correlations between average interval between heartbeats and perceived stress were consistently negative as well as the correlations between multiscale sample entropy and perceived stress, and multiscale fuzzy entropy and perceived stress, with some significant correlations as shown in Table 3. Also in Table 3, we listed correlation coefficients and corresponding $p$-values for low-frequency and perceived stress, showing inconsistent patterns.

We also explored correlations between the average level of stressor as calculated from the expert questionnaires and the four measures of HRV from the six phases as indicated earlier. The correlation coefficients were consistently negative for average interval between heartbeats, $r_s(8) = -.44$ to $-.21$ for performance 1 and $r_s(5) = -.47$ to $-.27$ for performance 2, with inconsistent patterns found for low-frequency and multiscale sample entropy and multiscale fuzzy entropy, though none were significant, $p > .05$.

**Discussion**

Our data suggest an effect of condition on HRV. Participants showed lower average interval between heartbeats, multiscale sample entropy, and multiscale fuzzy entropy during silent onstage periods compared to offstage periods. These HRV metrics were also lower for Performance 1, though not significantly so. On the whole, these data suggest that opera trainees have lower HRV and thus are more stressed during silent onstage periods compared to offstage periods and that they are potentially more stressed during Performance 1.

Low-frequency HRV was also less during silent periods. This occurred despite other HRV metrics indicating that silent periods are more stressful. Although the pattern of change in low-frequency HRV was contrary to general expectations, our results are consistent with previous research identifying lower low-frequency HRV during more stressful periods in performing musicians (Chanwimalueang et al., 2017). It should be noted that participants in our study were opera trainees, whereas Chanwimalueang et al. (2017) studied flutists and violinists. As mentioned in the introduction, performance conditions are often different between opera performers and instrumentalists.

Taken together, our results suggest that the first performance may be more stressful, and that it is more stressful to be onstage than it is to be backstage. The effect of performance on stress however may be restricted to offstage and silent onstage periods. When it comes to the actual singing periods, we found no indication of an effect of performance on HRV.

| Performance phase | Performance 1 | Performance 2 | $p$-values comparing Performance 1 to Performance 2 |
|-------------------|--------------|--------------|-----------------------------------------------|
| Before the performance | 6.11 (0.92) | 4.11 (0.81) | .018                                           |
| During the performance | 4.44 (1.05) | 2.78 (0.83) | .024                                           |
| $p$-values comparing before to during the performance | .027 | .023 |  |
Table 3. Correlation Between Perceived Stress Scale and Measures of HRV During Three Performance Conditions for Two Performances.

| HRV metric | Performance 1 | Performance 2 |
|------------|---------------|---------------|
|            | Offstage      | Silent        | Active        | Offstage | Silent | Active |
| AVNN       | \(r_{(7)} = -0.64, p = 0.063\) | \(r_{(7)} = -0.60, p = 0.088\) | \(r_{(7)} = -0.53, p = 0.141\) | \(r_{(4)} = -0.79, p = 0.059\) | \(r_{(4)} = -0.88, p = 0.020^*\) | \(r_{(4)} = -0.97, p = 0.001^*\) |
| LF         | \(r_{(7)} = 0.79, p = 0.012^*\) | \(r_{(7)} = 0.44, p = 0.235\) | \(r_{(7)} = 0.32, p = 0.396\) | \(r_{(4)} = -0.27, p = 0.612\) | \(r_{(4)} = 0.47, p = 0.348\) | \(r_{(4)} = 0.67, p = 0.142\) |
| MSE        | \(r_{(7)} = -0.84, p = 0.005^*\) | \(r_{(7)} = -0.06, p = 0.880\) | \(r_{(7)} = -0.55, p = 0.126\) | \(r_{(4)} = -0.74, p = 0.096\) | \(r_{(4)} = -0.44, p = 0.381\) | \(r_{(4)} = -0.18, p = 0.738\) |
| MFE        | \(r_{(7)} = -0.84, p = 0.005^*\) | \(r_{(7)} = 0.08, p = 0.829\) | \(r_{(7)} = -0.38, p = 0.313\) | \(r_{(4)} = -0.74, p = 0.096\) | \(r_{(4)} = -0.56, p = 0.249\) | \(r_{(4)} = -0.09, p = 0.868\) |

AVNN: average interval between heartbeats; LF: low-frequency; MFE: multiscale fuzzy entropy; MSE: multiscale sample entropy.

*significant effect.
The correlation between the experts’ ratings of the stress associated with each role and average interval between heartbeats was consistently negative, with inconsistent patterns for the correlation between the former and low-frequency, multiscale sample entropy, and multiscale fuzzy entropy, respectively. These mixed results could be explained by the small sample size of this study, but offer intriguing avenues for future research. Further research is needed to evaluate the relationship between the external evaluation of stress in different roles with HRV of performers.

Out of eight participants who indicated a difference between their stress levels onstage versus offstage, seven indicated they experience higher nervousness when they are offstage compared to onstage. This pattern is of particular interest because of the reversed pattern of results in the HRV data which suggested greater physiological stress onstage. Thus, opera trainees are either inaccurate in their self-assessment of stress or there are two different types of stress that are affecting performers.

While onstage opera trainees are confronted with high demands. Apart from the physiological demand of singing, performances take place under stage lighting and in costume, whose weight puts physical demands on performers in addition to the heat of stage lights. Additionally, monitoring signals from the conductor and colleagues onstage present cognitive demands. It is likely that decreased HRV onstage is due to these multiple demands. At the same time, it is possible that opera trainees perceive flow while they are performing onstage, reducing opportunities to actively reflect on the challenges of onstage performance.

Flow is a subjective state where individuals experience focus on the present moment, merging of action and awareness, loss of reflective self-consciousness, control of the situation, distortion of temporal experience, and intrinsic reward by the activity itself (Csikszentmihalyi, 1990). Challenges that match skills, clear proximal goals, and feedback on progress are conditions that promote flow experiences. While onstage, during a performance, opera trainees must intensely focus on their actions in the present moment. These conditions are conducive to flow (Nakamura & Csikszentmihalyi, 2014).

In contrast, waiting offstage between periods of onstage performance is not conducive to experiencing flow because the conditions described earlier are absent. Waiting is likely to free up cognitive capacity for reflecting on the challenging or perhaps even threatening characteristics of upcoming, onstage performing. Contingent on whether opera trainees think they are able to cope, negative stress might, therefore, be higher during these periods (Lazarus & Folkman, 1984). These disparate conditions for flow may explain the disconnect between HRV and self-reported measures of stress. Hence, HRV may be a more accurate measure of how physiologically stressful the situation in itself is, with demands onstage outweighing demands offstage.

In summary, we propose that the patterns in HRV that we observed reflect the physiological and cognitive demands of opera performance, with greater demands placed on the trainee while onstage. To this end, it will be of particular interest to further develop ways of assessing these demands through external ratings. We further propose that trainees’ perceived nervousness is lower onstage because of the experience of flow. In line with this proposal, there was a significant effect of the performance phase on the trainees’ self-perceived nervousness, such that they were more nervous before performances than during performances.

There was also a significant effect of performance, where self-reported nervousness of opera trainees was higher for Performance 1. This pattern is interesting in the context of the incongruent HRV data considering the offstage or onstage comparison, as greater self-reported nervousness for Performance 1 corresponded to lower HRV during Performance 1. Future work should study opera performance across more than two performances to determine if this pattern of decreased self-perceived stress continues.
Analysis of the relationships between HRV and other measures of self-report revealed a consistent negative association between HRV (as estimated by average interval between heartbeats, multiscale sample entropy, multiscale fuzzy entropy) and perceived stress. Here, higher perceived stress was associated with lower HRV. Together with the shown effects of performance condition on HRV itself, these results suggest that average interval between heartbeats, multiscale sample entropy, and multiscale fuzzy entropy can be used to observe stress in performers.

The association between low-frequency and perceived stress did not fall into this pattern and the results are inconsistent. Other studies showed increased psychological stress associated with an increase in low-frequency (Kim et al., 2018). However, given the reverse pattern both in our own data and in previous research (Chanwimalueang et al., 2017), we suggest that low-frequency may be an unreliable indicator of stress in performers. This finding adds to literature showing that frequency domain estimates of HRV may be unreliable during musical performance (Saboul et al., 2013). Thus, future research should consider using temporal and nonlinear analyses of HRV.

We found inconsistent patterns in the associations between our HRV variables and music performance anxiety. Two of the probed correlations were significant (multiscale sample entropy or multiscale fuzzy entropy and music performance anxiety), though only for one particular performance condition during one performance. Surprisingly, these two correlations were positive, indicating greater HRV with greater music performance anxiety. This is contrary to the expectation that higher music performance anxiety would lead to higher stress onstage (and thus, lower HRV). This could be explained by the low power of our analysis, but also by the fact that the Kenny Music Performance Anxiety Inventory was not validated for opera singers (Chang-Arana et al., 2018). The lack of significant correlations between the other measures of HRV and music performance anxiety is consistent with previous research reporting no such associations (van Fenema et al., 2017). This calls into question whether the Kenny Music Performance Anxiety Inventory captures how instrumentalists and singers deal with stress onstage.

One way of conceptualizing stress is outlined in the transactional model of stress by Lazarus and Folkman (1984). In the context of this model, the impact of any stressful event is moderated by how a person appraises it. Stress moderators are internal and external resources or vulnerabilities that modify how stress and its effects are experienced. We probed the association of music performance anxiety with a variety of these factors in order to explore the Kenny Music Performance Anxiety Inventory’s use as an indicator of performance stress.

Two potential internal resources are self-esteem and self-efficacy. Self-esteem is defined as “the awareness of good possessed by the self, and the level of global self-regard that one has for the self as a person” (Campbell & Fairey, 1985, p. 1108). Research has shown that it is a stable and universal concept (Altmann & Roth, 2018; Cai et al., 2009). Self-efficacy refers to “one’s overall beliefs or confidence in his/her ability to produce desired outcomes” (Bandura, 1977, p. 193). Both have been identified as important predictors of adjustment to stress (Lazarus & Folkman, 1984). They are inversely related to psychological indicators of stress and accordingly may decrease physiological responses to stress (Kivimäki & Kalimo, 1996) and predict music performance anxiety (Chan, 2011).

Similarly, some researchers argue that task mastery helps reduce performers’ anxiety (Matei & Ginsborg, 2017), although experience and practice do not eliminate the feelings of anxiety and stress in professional instrumentalists and singers (Nagel, 1990; Taborsky, 2007). There might be a link between age at which instrument training started and music performance anxiety: If instrument training started at age 7 or below, levels of music performance anxiety are lower (Zarza-Alzugaray et al., 2018). Generally, higher levels of music performance anxiety are reported for female instrumentalists (Iusca & Dafinoiu, 2012; Yondem, 2007), though again,
the Kenny Music Performance Anxiety Inventory has not previously been validated for opera singers specifically (Chang-Arana et al., 2018). The latter often also begin training much later than instrumentalists to accommodate bodily changes during puberty.

The correlation between the total music performance anxiety score and self-esteem and self-efficacy was negative, as expected, though not significantly so. Our data also indicate that an individual's music performance anxiety is not largely affected by their years of experience though our data suggest some association between music performance anxiety and other measures of experience, for example, the number of operas an individual has performed in. These results further call into question whether the Kenny Music Performance Anxiety Inventory can index stress as conceptualized by Lazarus and Folkman (1984). Furthermore, we found a negative correlation between music performance anxiety and negative emotionality. This stands in contrast to the positive correlation reported previously (Baumgartner & Schneider, 2020; Hong, 2010; Kendler et al., 2003; Lahey, 2009; Smith & Rickard, 2004).

This pattern may be due to generally higher levels of negative emotionality among opera singers (Sandgren, 2005). Furthermore, negative emotionality has been shown to be positively correlated with art interests and activities (Afhami & Mohammadi-Zarghan, 2018; Furnham & Chamorro-Premuzic, 2004). Thus, negative emotionality in opera singers may enhance their performance skills and help them adapt to different roles. Future research is needed to replicate this pattern in the relationship between negative emotionality and music performance anxiety in opera singers.

Our results have implications for stress management interventions in performers, in that they indicate that psychological stress is higher offstage. Stress in itself is not necessarily a negative thing for performers. However, stress can become negative if performance anxiety reaches clinical levels and impacts instrumentalists’ ability to perform (Gembris et al., 2018; Nagel, 2018). For example, two of our participants’ responses on the Kenny Music Performance Anxiety Inventory suggested potentially performance hindering levels of anxiety (Kenny, 2015).

**Limitations**

One major limitation to our study is the small sample size. We recruited from a small overall population (i.e., opera trainees who were cast). While this limits our ability to draw conclusions about associations between variables in this dataset, our ecologically valid dataset provides an impetus to study these associations in larger samples in the future. It should also be noted that the study of stress in opera trainees is novel. Hence, our study extends the work on musicians’ stress which has been done primarily with instrumentalist participants. We would also like to caution against applying our findings to instrumentalists. As mentioned in the introduction, performance conditions often differ between opera performers and instrumentalists.

The artificial nature of performance experiments can decrease ecological validity. In Chanwimalueang et al. (2017), for example, measurements were collected in front of a three-person audience, which is more representative of an audition than a real-world music performance. The presence of a live audience may have a significant impact on performers’ stress levels. Therefore, our results add to the literature on performance stress by measuring HRV in a context with high ecological validity.

At the same time, high ecological validity makes it difficult to include experimental control groups. In our case, most participants performed different roles in the opera, making it difficult to equate and compare measurements of performers. Different roles have different performance demands such as time spent onstage, and material of varying difficulty, though our exploratory analyses showed no significant association between the latter with HRV. In future iterations, we plan to collect data from a whole double cast so that each role is measured on two performers.
Another unknown factor in our study is what participants actually did when they were off-stage, as we were only able to videotape activities onstage during the performance. While our assumption is that silent periods onstage are most comparable to offstage periods, participants could theoretically have spent time offstage singing. Future research may attempt to better regularize what happens during the latter periods.

**Conclusion**

Our study found that HRV in opera trainees is lower in the silent onstage periods compared to offstage periods. These data suggest that physiological stress is greater onstage than offstage. This pattern is the reverse of what was noted for self-reported nervousness. This suggests that there are two types of stress that are relevant for opera performance: psychological stress, which is felt more keenly offstage, and physiological stress, which is greater onstage. Future research may wish to expand on this finding by studying professional opera performers and instrumentalists.

**Acknowledgements**

The authors would like to acknowledge the invaluable help of Julie Youngers and Allison Anderson in collecting data, Dr Janet Werker for insightful discussion and feedback, and Dr Theerasak Chanwimalueang and his co-authors for sharing their algorithms.

**Funding**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a grant from the Peter Wall Institute for Advanced Studies. Anja-Xiaoxing Cui is supported in part by funding from the Social Sciences and Humanities Research Council of Canada.

**ORCID iD**

Anja-Xiaoxing Cui [https://orcid.org/0000-0002-5003-6867](https://orcid.org/0000-0002-5003-6867)

**Supplemental material**

Supplemental material for this article is available online.

**References**

Afhami, R., & Mohammadi-Zarghan, S. (2018). The big five, aesthetic judgment styles, and art interest. *Europe’s Journal of Psychology, 14*(4), 764–775. [https://doi.org/10.5964/ejop.v14i4.1479](https://doi.org/10.5964/ejop.v14i4.1479)

Altmann, T., & Roth, M. (2018). The self-esteem stability scale (SESS) for cross-sectional direct assessment of self-esteem stability. *Frontiers in Psychology, 9*, 91. [https://doi.org/10.3389/fpsyg.2018.00091](https://doi.org/10.3389/fpsyg.2018.00091)

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. [https://doi.org/10.1037//0033-295X.84.2.191](https://doi.org/10.1037//0033-295X.84.2.191)

Baumgartner, J. N., & Schneider, T. R. (2020). Personality and stress. In *Encyclopedia of personality and individual differences* (pp. 3699–3704). [https://doi.org/10.1007/978-3-319-28099-8_2115-1](https://doi.org/10.1007/978-3-319-28099-8_2115-1)

Bernardi, L., Wdowczyk-Szulc, J., Valenti, C., Castoldi, S., Passino, C., Spadacini, G., & Sleight, P. (2000). Effects of controlled breathing, mental activity and mental stress with or without verbalization on heart rate variability. *Journal of the American College of Cardiology, 35*(6), 1462–1469. [https://doi.org/10.1016/s0735-1097(00)00595-7](https://doi.org/10.1016/s0735-1097(00)00595-7)

Burin, A. B., & Osorio, F. L. (2017). Music performance anxiety: A critical review of etiological aspects, perceived causes, coping strategies and treatment. *Revista De Psiquiatria Clinica, 44*(5), 127–133. [https://doi.org/10.1590/0101-60830000000136](https://doi.org/10.1590/0101-60830000000136)
Cai, H., Wu, Q., & Brown, J. D. (2009). Is self-esteem a universal need? Evidence from the People’s Republic of China. *Asian Journal of Social Psychology, 12*(2), 104–120. https://doi.org/10.1111/j.1467-839X.2009.01278.x

Campbell, J. D., & Fairey, P. J. (1985). Effects of self-esteem, hypothetical explanations, and verbalization of expectancies on future performance. *Journal of Personality and Social Psychology, 48*(5), 1097–1111. https://doi.org/10.1037/0022-3514.48.5.1097

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

Chan, M. Y. (2011). The relationship between music performance anxiety, age, self-esteem, and performance outcomes in Hong Kong music students [Doctoral dissertation, Durham University]. http://etheses.dur.ac.uk/637/

Chang-Arana, Á. M., Kenny, D. T., & Burga-León, A. A. (2018). Validation of the Kenny music performance anxiety inventory (K-MPAI): A cross-cultural confirmation of its factorial structure. *Psychology of Music, 46*(4), 551–567. https://doi.org/10.1177/0305735617717618

Chanimalueang, T., Aufegger, L., Adjei, T., Wasley, D., Cruder, C., Mandic, D. P., & Williamon, A. (2017). Stage call: Cardiovascular reactivity to audition stress in musicians. *PLoS ONE, 12*(4), Article e0176023. https://doi.org/10.1371/journal.pone.0176023

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a new general self-efficacy scale. *Organizational Research Methods, 4*(1), 62–83. https://doi.org/10.1177/109442810141004

Chen, W., Wang, Z., Xie, H., & Yu, W. (2007). Characterization of surface EMG signal based on fuzzy entropy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 15*(2), 266–272. https://doi.org/10.1109/TNSRE.2007.897025

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior, 24*(4), 385–396. https://doi.org/10.2307/2136404

Costa, M., Goldberger, A. L., & Peng, C. K. (2002). Multiscale entropy analysis of complex physiologic time series. *Physical Review Letters, 89*(6), 068102. https://doi.org/10.1103/physrevlett.89.068102

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653–665. https://doi.org/10.1016/0191-8869(92)90236-I

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience* (1st ed.). Harper & Row.

Delaney, J. P., & Brodie, D. A. (2000). Effects of short-term psychological stress on the time and frequency domains of heart-rate variability. *Perceptual and Motor Skills, 91*(2), 515–524. https://doi.org/10.2466/pms.2000.91.2.515

Dimsdale, J. E. (2008). Psychological stress and cardiovascular disease. *Journal of the American College of Cardiology, 51*(13), 1237–1246. https://doi.org/10.1016/j.jacc.2007.12.024

Dishman, R. K., Nakamura, Y., Garcia, M. E., Thompson, R. W., Dunn, A. L., & Blair, S. N. (2000). Heart rate variability, trait anxiety, and perceived stress among physically fit men and women. *International Journal of Psychophysiology, 37*(2), 121–133. https://doi.org/10.1016/s0167-8760(00)00085-4

Elite, HRV. (2020, August 6). How do you calculate the HRV score? Elite HRV. https://help.elitehrv.com/article/54-how-do-you-calculate-the-hrv-score

Furnham, A., & Chamorro-Premuzic, T. (2004). Personality, intelligence, and art. *Personality and Individual Differences, 36*(3), 705–715. https://doi.org/10.1016/s0191-8869(03)00128-4

García-González, M. A., Vázquez-Seisdedos, C., & Pallás-Areny, R. (2000). Variations in breathing patterns increase low frequency contents in HRV spectra. *Physiological Measurement, 21*(3), 417–423. https://doi.org/10.1088/0967-3334/21/3/307

Gembris, H., Heye, A., & Seifert, A. (2018). Health problems of orchestral musicians from a life-span perspective: Results of a large-scale study. *Music & Science, 1*, 205920431773980. https://doi.org/10.1177/2059204317739801

Gidron, Y. (2013). Trait anxiety. In M. D. Gellman & J. R. Turner (Eds.), *Encyclopedia of behavioral medicine*. Springer. https://doi.org/10.1007/978-1-4419-1005-9

Gilgen-Ammann, R., Schweizer, T., & Wyss, T. (2019). RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. *European Journal of Applied Physiology, 119*(7), 1525–1532. https://doi.org/10.1007/s00421-019-04142-5
Harmat, L., & Theorell, T. (2010). Heart rate variability during singing and flute playing. *Music and Medicine, 2*(1), 10–17. https://doi.org/10.1177/1943862109354598

Hong, R. Y. (2010). Neuroticism, anxiety sensitivity thoughts, and anxiety symptomatology: Insights from an experience-sampling approach. *Cognitive Therapy and Research, 34*(3), 254–262. https://doi.org/10.1007/s10608-009-9231-4

Humara, M. (1999). The relationship between anxiety and performance: A cognitive–behavioral perspective. *Athletic Insight: The Online Journal of Sport Psychology, 1*(2). https://psycnet.apa.org/record/2001-03339-001

Iusca, D., & Dafinoiu, I. (2012). Performance anxiety and musical level of undergraduate students in exam situations: The role of gender and musical instrument. *Procedia, Social and Behavioral Sciences, 33*, 448–452. https://doi.org/10.1016/j.sbspro.2012.01.161

Kendler, K. S., Gardner, C. O., & Prescott, C. A. (2003). Personality and the experience of environmental adversity. *Psychological Medicine, 33*(7), 1193–1202. https://doi.org/10.1017/S0033291703008298

Kennedy, B. L., Schwab, J. J., Morris, R. L., & Beldia, G. (2001). Assessment of state and trait anxiety in subjects with anxiety and depressive disorders. *Psychiatric Quarterly, 72*(3), 263–276. https://doi.org/10.1023/A:1010305200087

Kenny, D. T. (2011). *The psychology of music performance anxiety*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199586141.001.0001

Kenny, D. T. (2015). Identifying cut-off scores for clinical purposes for the Kenny Music Performance Anxiety Inventory (K-MPAI) in a population of professional orchestral musicians. *Polish Psychological Bulletin* (Invited paper for the special edition on music performance anxiety). https://doi.org/10.13140/RG.2.1.2481.2244

Kenny, D. T., Davis, P., & Oates, J. (2004). Music performance anxiety and occupational stress amongst opera chorus artists and their relationship with state and trait anxiety and perfectionism. *Journal of Anxiety Disorders, 18*(6), 757–777. https://doi.org/10.1016/j.janxdis.2003.09.004

Kim, H., Cheon, E., Bai, D., Lee, Y. H., & Koo, B. (2018). Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investigation, 15*(3), 235–245. https://doi.org/10.30773/pi.2018.07.17

Kivimäki, M., & Kalimo, R. (1996). Self-esteem and the occupational stress process: Testing two alternative models in a sample of blue-collar workers. *Journal of Occupational Health Psychology, 1*(2), 187–196. https://doi.org/10.1037/1076-8998.1.2.187

Ladd, S. L., & Gabrieli, J. D. E. (2015). Trait and state anxiety reduce the mere exposure effect. *Frontiers in Psychology, 6*. 701. https://doi.org/10.3389/fpsyg.2015.00701

Lahey, B. B. (2009). Public health significance of neuroticism. *The American Psychologist, 64*(4), 241–256. https://doi.org/10.1037/a0015309

Lazarus, R. S., & Folkman, S. (1984). *Stress, appraisal, and coping*. Springer. https://doi.org/10.1007/978-1-4419-1005-9_215

Lee, E. (2012). Review of the psychometric evidence of the perceived stress scale. *Asian Nursing Research, 6*(4), 121–127. https://doi.org/10.1016/j.anr.2012.08.004

Lehmann, A. C., Sloboda, J. A., & Woody, R. H. (2007). *Psychology for musicians: Understanding and acquiring the skills*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195146103.001.0001

Lipsitz, L. A., & Goldberger, A. L. (1992). Loss of “complexity” and aging: Potential applications of fractals and chaos theory to senescence. *Journal of the American Medical Association, 267*(13), 1806–1809. https://doi.org/10.1001/jama.1992.03480130122036

Matei, R., & Ginsborg, J. (2017). Music performance anxiety in classical musicians—What we know about what works. *Bjpsych International, 14*(2), 33–35. https://doi.org/10.1192/B205647400000174

Müller, A. M., Wang, N. X., Yao, J., Tan, C. S., Low, I. C. C., Lim, N., Tan, J., Müller-Riemenschneider, F., & Tan, A. (2019). Heart rate measures from wrist-worn activity trackers in a laboratory and free-living setting: Validation study. *JMIR Mhealth and Uhealth, 7*(10), e14120. https://doi.org/10.2196/14120

Nagel, J. J. (1990). Performance anxiety and the performing musician: A fear of failure or a fear of success. *Medical Problems of Performing Artists, 5*(1), 37–40.
Wells, R., Outhred, T., Heathers, J. A. J., Quintana, D. S., & Kemp, A. H. (2012). Matter over mind: A randomised-controlled trial of single-session biofeedback training on performance anxiety and heart rate variability in musicians. *PLoS ONE, 7*(10), Article e46597. https://doi.org/10.1371/journal.pone.0046597

Williamon, A., Aulegger, L., Wasley, D., Looney, D., & Mandic, D. P. (2013). Complexity of physiological responses decreases in high-stress musical performance. *Journal of the Royal Society Interface, 10*(89), 20130719. https://doi.org/10.1098/rsif.2013.0719

Xiong, G., Zhang, L., Liu, H., Zou, H., & Guo, W. (2010). A comparative study on ApEn, SampEn and their fuzzy counterparts in a multiscale framework for feature extraction. *Journal of Zhejiang University SCIENCE A, 11*(4), 270–279. https://doi.org/10.1631/jzus.A0900360

Yondem, Z. D. (2007). Performance anxiety, dysfunctional attitudes and gender in university music students. *Social Behavior and Personality: An International Journal, 35*(10), 1415–1426. https://doi.org/10.2224/sbp.2007.35.10.1415

Zarza-Alzugaray, F. J., Orejudo, S., Casanova, O., & Aparicio-Moreno, L. (2018). Music performance anxiety in adolescence and early adulthood: Its relation with the age of onset in musical training. *Psychology of Music, 46*(1), 18–32. https://doi.org/10.1177/0305735617691592