Complexity of physiological responses decreases in high-stress musical performance

Aaron Williamon, Lisa Aufegger, David Wasley, David Looney and Danilo P. Mandic

1Centre for Performance Science, Royal College of Music, Prince Consort Road, London SW7 2BS, UK
2Cardiff School of Sport, Cardiff Metropolitan University, Cyncoed Campus, Cyncoed Road, Cardiff CF23 6XO, UK
3Department of Electrical and Electronic Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

For musicians, performing in front of an audience can cause considerable apprehension; indeed, performance anxiety is felt throughout the profession, with wide ranging symptoms arising irrespective of age, skill level and amount of practice. A key indicator of stress is frequency-specific fluctuations in the dynamics of heart rate known as heart rate variability (HRV). Recent developments in sensor technology have made possible the measurement of physiological parameters reflecting HRV non-invasively and outside of the laboratory, opening research avenues for real-time performer feedback to help improve stress management. However, the study of stress using standard algorithms has led to conflicting and inconsistent results. Here, we present an innovative and rigorous approach which combines: (i) a controlled and repeatable experiment in which the physiological response of an expert musician was evaluated in a low-stress performance and a high-stress recital for an audience of 400 people, (ii) a piece of music with varying physical and cognitive demands, and (iii) dynamic stress level assessment with standard and state-of-the-art HRV analysis algorithms such as those within the domain of complexity science which account for higher order stress signatures. We show that this offers new scope for interpreting the autonomic nervous system response to stress in real-world scenarios, with the evolution of stress levels being consistent with the difficulty of the music being played, superimposed on the stress caused by performing in front of an audience. For an emerging class of algorithms that can analyse HRV independent of absolute data scaling, it is shown that complexity science performs a more accurate assessment of average stress levels, thus providing greater insight into the degree of physiological change experienced by musicians when performing in public.

1. Introduction

Performing music in public requires the management of intense physical and mental demands. How musicians perceive and respond to these demands, and deliver high-quality performances consistently under pressure, can determine not only the success of single events but also the path and length of their careers [1,2].

In this respect, musicians are not unlike elite performers in other domains. Under intense stress, physiological and psychological responses such as heart rate and level of state anxiety are markedly increased for both those who must work hard physically, such as athletes [3], as well as those whose work requires mental exertion, such as surgeons [4] and chess grandmasters [5]. While the analysis of physiological responses is well explored in sports science and in many clinical fields, studies in music, particularly those examining stress in real-world contexts and at the highest of international levels, are rare.

Stress is managed by the autonomic nervous system (ANS). In particular, a reaction to stress can be characterized by the interactions between two ANS...
components: the parasympathetic nervous system (PNS), associated with homeostasis and balance, and the sympathetic nervous system (SNS), associated with greater arousal [6]. SNS/PNS interactions have been shown to influence the temporal fluctuations of the peak-to-peak times in the electrocardiogram (ECG)—the R-to-R (RR) interval—known as heart rate variability (HRV) [7,8]. The simplest measures of HRV are the mean of the RR time series and its standard deviation about the mean. Stress often causes a decrease in the mean RR (increased heart rate) in healthy individuals, with the opposite effect in chronically stressed individuals [9], while in the case of standard deviation, certain studies have found that it fails to vary significantly in conditions of mental stress [10]. Moreover, both the mean and standard deviation statistics are based on the absolute magnitude of the RR interval, whereas in many applications relative measures of a process are understood to exhibit a greater consistency for inter-individual comparisons. For instance, the resting heart rate can vary considerably from individual to individual.

In recent years, studies of HRV in the stress assessment context have focused on its low-frequency (LF) and high-frequency (HF) components: 0.04–0.15 Hz and 0.15–0.4 Hz, respectively (for a review, see [6–8]). It has been widely accepted that the HF element reflects PNS activity whereas the LF element, although more complicated, reflects SNS activity [11]. Changes induced by pharmacological stress have been shown to cause an increase in SNS activity while reciprocally causing a withdrawal in PNS activity, a phenomenon known as the sympathetic-vagal balance. However, in real-world contexts, the response of the ANS to stress is diverse and depends on the nature of the stressor (physical/psychological) and in some cases, the individual [6]. For instance, a study of psychological stress [12] found that the extent of changes in PNS activity relative to SNS activity is person dependent. Nonetheless, it has been proposed that the LF/HF power ratio, a relative measure, can characterize the ‘balance’ relationship between the SNS and PNS [13,14] and has been widely used to study the effect of stress on performance.

Nakahara et al. [15] compared the LF/HF ratio elicited while musicians performed and listened to music, finding a higher ratio during performance. Harmat & Theorell [16] studied HR and HRV in professional singers and flautists during low- and high-stress performances. They found increased HR and suppressed HRV in the high-stress condition, but contrary to findings from previous research [6], LF power was significantly lower in high stress. Harmat et al. [17] examined HR and HRV in expert pianists while playing a familiar piece and while sight-reading a technically demanding unfamiliar piece. They found significantly higher LF power in the latter condition, which corresponded to a more cognitively demanding and (by implication) more stressful task.

The ambiguities and inconsistencies encountered using the LF/HF model might be explained by recent work [9,11], which argued that it oversimplifies the complex relationship between the SNS and PNS and challenged its accuracy. This has motivated us to investigate more appropriate methods for modelling the variable interactions within RR rhythms in conditions of stress.

Complexity science quantifies the ability of a living system to adapt to changes in the environment characterized by long-term auto- and cross-correlations within its physiological responses (coupled dynamics) at different scales. Multiscale sample entropy (MSE) is one such method that evaluates signal regularity, determined by sample entropy (SE), across multiple temporal scales and is particularly suited to revealing long-range correlations—a key feature of complex systems [18,19]. According to MSE theory, a complex system exhibits high sample entropy at multiple temporal scales. The method has been used in numerous human-centred applications and has demonstrated a good agreement with complexity-loss theory, which asserts that the complexity of physiological time series of an organism under constraints (owing to illness, ageing or stress) is lower than for unconstrained (healthy) organisms [20–22]. There are several advantages of the MSE method in HRV analysis. The algorithm can account for nonlinear couplings, enabling greater accuracies compared with standard linear measures (mean, standard deviation and LF/HF ratio), and is a relative measure (examines RR fluctuations independent of their absolute magnitude) making it more suitable for inter-individual analysis but without making rigid assumptions about the underlying generating mechanisms as is the case with other relative measures (LF/HF ratio). Studies of stress level changes induced in HRV by physical exertion [23,24] or meditation [25] suggest that complexity—determined by the MSE approach—is lowest during states of high stress, a result that is consistent with the complexity-loss theory. Despite the potential of the method in the study of stress, the precise data conditioning and pre-processing steps undertaken prior to MSE analysis are often not reported, yet these have a major impact on the coherence and interpretation of the results.

1.1. Aims of this study
Musical performance requires considerable motor precision integrated with sustained management of cognitive, perceptual and social processes and is a natural domain for studying the response to high-stress performance situations [26]. While consistency in executing domain-specific skills over time is a characteristic of expertise in any domain [27,28], the physical control exhibited by an expert classical musician in repeated performances [29,30] offers a unique opportunity to investigate the degree of stress caused by public performance.

In our study, the HRV of a concert pianist was assessed for performances of the same piece in low-stress and high-stress conditions. We set out to dynamically examine stress signatures caused by (i) varying physical and cognitive demands within the musical piece (identical across performances) and (ii) audience-induced anxiety (different across performances). For rigour, HRV analysis was performed using both standard and state-of-the-art techniques with identical pre-processing (considered frequency range, identical time windows) applied where relevant, in this way ensuring a fair comparison between the analysis methods.

2. Material and methods
2.1. Participant
Melvyn Tan (born 1956) is an internationally renowned pianist and performs regularly in many of the world’s leading concert halls.

2.2. Procedure
A preliminary health screening was first conducted. ECG data were recorded for performances in: (i) a low-stress condition, where only the performer and research team were present, and (ii) a high-stress condition with an audience of 400 people at
the 2012 Cheltenham Music Festival. The data were collected using a wireless Zephyr Bioharness [31,32] at a sampling rate of 250 Hz. Analysis was focused on data obtained during the first piece in the recital programme, J. S. Bach’s English Suite in A minor (BWV 807), where early stages of performance are particularly physically and psychologically stressful [33,34].

2.3. Data treatment

Each performance produced approximately 20 min of ECG data. The time difference between successive R peaks in ECG was estimated, which was converted into an RR time series using cubic spline interpolation with samples at regular time intervals of 0.25 s. The RR signal was bandpass filtered (0.04–0.4 Hz) via a fourth-order Butterworth filter before estimating the following features using overlapping windows of the same length:

- standard deviation of the RR signal about its mean;
- power in the LF (0.04–0.15 Hz) and HF components (0.15–0.4 Hz) of the RR signal obtained using a fourth-order Butterworth filter;
- LF/HF ratio obtained from the estimated power in the LF and HF bands; and
- SE estimated at different timescales for the complete frequency range (0.04–0.4 Hz), the LF range (0.04–0.15 Hz) and the HF range (0.15–0.4 Hz). In all cases, each windowed segment was normalized (zero mean, unit variance) before estimating the SE with the embedding dimension and tolerance level at 2 and 0.15, respectively (see Multiscale sample entropy below).

Windows of 7 min length were selected as the longest period of the considered RR component was 25 s (0.04 Hz), and at least 10 times the lowest oscillation period is advised in HRV analysis to sample short-term variations adequately [7]. It is worth noting that the LF/HF ratio and the MSE method, when estimated over normalized data segments, are relative measures, and do not depend on the absolute scaling of the RR data.

2.4. Multiscale sample entropy

MSE estimation is performed by two steps:

- The different temporal scales are estimated by coarse graining (moving average) the N-sample time series, \( \{ x_i \}, i = 1, \ldots, N \). For a scale factor, \( e \), the corresponding coarse-grained time series is given by: \( x'_i = (1/e) \sum_{j=1}^{e} x_{i+je}, \) for \( i = (j-1)e + 1, \ldots, je \) and \( j = 1, \ldots, N/e \).
- The SE is evaluated for each intrinsic scale \( x'_i \). Underpinning the method is the estimation of the conditional probability that two similar sequences will remain similar when the next data point is included. First, composite delay vectors of the scale are formed, with embedding dimension \( M \), and the average number of neighbouring delay vectors for a given tolerance level, \( r \), is estimated. This is known as the frequency-of-occurrence and reflects the level of self-similarity within the scale. This process is repeated for an embedding dimension of \( M + 1 \), and the ratio of the two frequency-of-occurrence values gives the SE of the scale.

For further information on the MSE method, see Costa et al. [18,19].

3. Results

Table 1 shows the completion time of each movement within the piece of music. The time difference between the two performances was 24 s, reflecting a high degree of consistency over a 20 min task and enabling a fair comparison across performances. The performer reported an increase in perceived pressure during the performance in front of the audience. He furthermore reported that the Prelude and Courante (the first and third movements) were the most challenging in both scenarios.

To provide insight into the level of resolution in time afforded by the standard deviation, LF/HF ratio and MSE methods, some of the movement ending times are shown in figure 1 relative to the window length (7 min). Figure 2 shows the results of the basic measures of HRV. Figure 2a shows the RR interval time series for the same performance under the low-stress (grey line) and high-stress (black line) performance conditions. The mean RR interval for the high-stress condition was significantly lower than that for the low-stress condition (the Bhattacharyya coefficient, a measure of the amount of overlap between two distributions, was zero). Figure 2b shows the standard deviation of the RR time series filtered within the frequency range 0.04–0.4 Hz. The standard deviation was lower for the high-stress performance (black line), conforming with some studies of the effects of stress on RR standard deviation [9]. There was a high level of similarity between the relative changes in standard deviation for each of the two performances: the standard deviation decreased at around 600 s. This result may support the reported difficulty experienced during the Courante, which ended at 578 s (high-stress performance), and indicates a reduction in stress possibly caused by the relief at having completed and passed through the most challenging parts of the piece, the reduced physical demands of the subsequent musical material, or both.

Figure 3 shows the total power in the LF bands (figure 3a(iii)) and HF bands (figure 3a(ii)). The high-stress performance (black line) resulted in reduced HF activity but also in a decrease in the LF activity (grey line). Thus the LF/HF ratio, as shown in figure 3b, was as expected for the initial stages of performance—it was highest for the high-stress condition—but for performance times after approximately 600 s, the ratio of the low-stress condition was highest. This runs counter to predictions of the physiological stress model based on an increase in the LF/HF ratio [6] and suggests that the ratio is inconsistent. Nonetheless, the relative decreases observed for the standard deviation features after approximately 600 s, potentially caused by the shift into less challenging movements of the piece, is also found in the LF/HF analysis.

<table>
<thead>
<tr>
<th>Movement</th>
<th>Low-stress performance (s)</th>
<th>High-stress performance (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prelude</td>
<td>255</td>
<td>259</td>
</tr>
<tr>
<td>Allemende</td>
<td>462</td>
<td>485</td>
</tr>
<tr>
<td>Courante</td>
<td>563</td>
<td>578</td>
</tr>
<tr>
<td>Sarabande</td>
<td>802</td>
<td>814</td>
</tr>
<tr>
<td>Bourrée I and II</td>
<td>1040</td>
<td>1048</td>
</tr>
<tr>
<td>Gigue</td>
<td>1236</td>
<td>1212</td>
</tr>
</tbody>
</table>

*The performer reported that the Prelude and Courante were the most challenging movements of the piece.*
Figure 1. The vertical bars denote times that the performer completed the first four movements: (1) Prelude, (2) Allemande, (3) Courante and (4) Sarabande. Grey bars denote the end times for the low-stress performance, and black bars the end times for the high-stress performance. The first and third movements, denoted by asterisk (*), were reported as being the most challenging. The horizontal arrow represents the length of the window used in the standard deviation, LF/HF ratio and MSE analyses, providing some insight into the level of time resolution afforded by the methods. (Online version in colour.)

Figure 2. (a) The RR interval time series. (b) The standard deviation of the RR time series (bandpass filtered within the range 0.04–0.4 Hz). In all figures, grey lines denote the low-stress performance, and black lines the high-stress performance.

Figure 3c,d shows the results of the SE analysis applied to the complete frequency range (0.04–0.4 Hz) for the first (no coarse graining) and second scale factors, respectively. The results indicate lower complexity, particularly at the second scale factor, for the high-stress performance (black line), which is in agreement with previous results: high-stress conditions yield lower complexity. In this way, the results are similar to those for the LF/HF ratio (figure 3b), and yet a greater separation between the two performances was facilitated by the MSE analysis (figure 3d). Also, a relative increase in complexity was found in both the high- and low-stress performances at around 600 s, which is consistent with the standard deviation and LF/HF analyses. For the same scale factors, the SE of the LF band only is given in figure 3e,f, and the SE for the HF band only is given in figure 3g,h. The SE analysis for the HF band discriminates between the high-stress (black line) and low-stress (grey line) performances and is consistent with complexity-loss theory, but there was no separation for the LF band. In both cases (figure 3e,h), however, a relative increase in complexity was not observed at around 600 s.

4. Discussion
This study presents the first rigorous examination of real-world autonomic response in the musical performance domain, demonstrating the degree of physiological change experienced by an expert musician when performing in public. Both standard and state-of-the-art tools were used to examine dynamically the components of heart rate governed by autonomic control, revealing signatures in HRV that indicate higher stress levels caused by public performance and also technically challenging movements in the considered piece.

In this instance, basic measures (mean RR and RR standard deviation) were able to distinguish clearly between the states of low and high stress. The results of the standard deviation analysis also provide insight into the reported difficulties experienced by the performer during the first and third movements of the piece, as the relative values decreased once the third movement ended, suggesting a relative decrease in stress. It is well known, however, that such basic measures are not always reliable [9,10]. Another disadvantage is that these measures are based on the absolute magnitude of the RR signal whereas, in general, relative measures of a process are understood to exhibit greater consistency.

The results of the two relative measures, the LF/HF ratio and the MSE method, were similar when applied to the same RR frequency components: both methods indicate a higher level of stress for the first part of the performance and a relative decrease in stress after the end of the third movement. However, the LF/HF ratio exhibits inconsistent results from 600 s suggesting the performance for the audience became less stressful than the performance without an audience, contradicting the reported experience of the musician. On the other hand, the MSE method clearly shows lower regularity across the intrinsic data scales for the state of high stress—indicating a lower complexity of the physiological state. This is in agreement with the complexity-loss theory for living organisms under constraint: stress causes a reduction in complexity. In addition, it was found that the complexity of the HF band, typically associated with PNS activity, was
reduced for the state of high stress. Little change in complexity was observed in the LF band.

While this case study offers valuable insight into how MSE analysis can be applied to data collected in real-world contexts, subsequent investigations are needed to establish the extent in which the results generalize to larger samples of expert performers and to those at lower levels of skill. Also, the MSE method should be extended to investigate jointly the dynamics of other physiological parameters (e.g. respiration rate) under stressful conditions. Finally, the utility of physiological complexity as a measure in stress-reduction interventions needs to be investigated, such as in cognitive-behavioural training or biofeedback. The findings reported here also offer promising new avenues for identifying stress responses in a wide range of performance situations, both in music and in other performance domains.

This study was conducted according to ethical guidelines of the British Psychological Society. Our research is a case study in which the participant is named and large amounts of personal (physiological) data were collected. For ethical reasons, we are unable to submit our dataset to a public repository. We would, however, consider releasing extracts of the data to third parties upon request (e.g. for

---

**Figure 3.** (a) The power in the LF band of the RR time series (i) and the HF band (ii). (b) The LF/HF ratio of the pre-processed (filtered, normalized) RR time series. (c) The SE for the first scale factor of the normalized time series, bandpass filtered within the range 0.04–0.4 Hz. (d) The SE for the second scale factor of the normalized time series, bandpass filtered within the range 0.04–0.4 Hz. (e) The SE for the first scale factor of the normalized time series, bandpass filtered within the range 0.04–0.15 Hz. (f) The SE for the second scale factor of the normalized time series, bandpass filtered within the range 0.04–0.15 Hz. (g) The SE for the first scale factor of the normalized time series, bandpass filtered within the range 0.15–0.4 Hz. (h) The SE for the second scale factor of the normalized time series, bandpass filtered within the range 0.15–0.4 Hz. In all figures, grey lines denote the low-stress performance, and black lines the high-stress performance.
verification of calculations), but only after obtaining written permission from the participant.

Acknowledgements. We are grateful to Melvyn Tan for generously contributing his time and musical prowess to the project. We also thank Meurig Bowen and Penny Wright for initial encouragement to undertake the research and help incorporating our data collection into the busy schedule of the 2012 Cheltenham Music Festival. All authors contributed extensively to the work presented in this paper.

Funding statement. The authors declare no competing financial interests.

References