# scientific reports

OPEN

# Randomized cross-over trial comparing stress responses amongst undergraduates and surgeons with and without background music during simulated surgery

Anantha Narayanan[1,2,3✉], Manar Khashram[1,2] & James P. Fisher[3]

High stress may diminish a surgeon's performance in the operating room. Music is perceived to reduce stress in this setting, however the psycho-physiological effects of music on intra-operative stress in inexperienced and experienced operators is incompletely understood. The effect of music on the psychological (Six-Item State-Trait Anxiety Inventory [STAI-6] and Surgical Taskload Index [SURG-TLX]) and physiological responses (e.g., heart rate variability) was determined to a simulated surgical task (carotid patch-angioplasty) in 15 medical students and 12 vascular surgeons under stressing conditions in a randomised crossover design. Music did not affect the speed or accuracy of the simulated surgical stress task performance. While the surgical task increased SURG-TLX scores from baseline to control (Δ32 [22–42]; mean difference [95% confidence interval]) and to music (Δ30 [20–40]), and increased STAI-6 scores in both conditions, there was no difference between music and control. The surgical task also increased heart rate (peak Δ5.1 bpm [3.0–7.1] vs. baseline $p < 0.0001$) and an index of cardiac sympathetic nervous system activity, and reduced an index of parasympathetic nervous system activity, with the latter two exacerbated by music (SNS index: 0.14 [0.004–0.27], $p = 0.042$; PNS index: − 0.11 [− 0.22 to − 0.008], $p = 0.032$). The more experienced group performed faster and more accurately than the inexperienced group, but there were no psychological or physiological differences in their responses to music. Despite previous research identifying generally positive surgeon perceptions of music on the intra-operative experience of stress, herein, background music failed to improve surgical task performance or attenuate subjective ratings of task load and anxiety, and physiological arousal.

It is common for surgeons to experience stress while they are performing surgery[1]. High levels of physiological and/or psychological stress have the potential to interfere with the expert execution of technical and non-technical skills (e.g., communication, teamwork and rapid decision-making) in a stressful environment[2,3]. The flow-on effects of intra-operative stress may have the potential to impair surgical performance and thereby worsen patient outcome[2,3]. Furthermore, acute stress phenotype may contribute to a burnout / chronic stress phenotype[4]. Therefore, the application of stress mitigation strategies within the operating room (OR) may be important to ensure the ongoing delivery of excellent surgical care through the execution of technical and non-technical skills[1–3].

Music is regularly played in ORs throughout the world and is a common way in which surgeons can alter their operating environment[5–9]. Many surgical clinicians find music to be a generally favorable part of the theatre environment[7–12], with music being seen as improving calmness[5], stress[13,14], mood[8], surgeon and overall team performance[7], however concerns remain over its distracting properties and risk to disrupt communication[7,11,13,15,16]. In order to investigate the effect of environment on a surgeon's experience of stress and task performanc, a surgical simulation task was used, as this permits superior control of confounding variables

[1]Department of Surgery, Faculty of Medicine and Health Sciences, University of Auckland, Auckland, New Zealand. [2]Department of Vascular Surgery and Endovascular Surgery Waikato Hospital, Waikato, New Zealand. [3]Department of Physiology, Faculty of Medicine and Health Sciences, University of Auckland, Auckland, New Zealand. ✉email: anantha_narayanan@hotmail.com

and more extensive physiological monitoring than in the real-world[17]. For example, Oomens et al. observed that music improved task performance, instrument handling efficiency and reduced SURG-TLX parameters during a simple surgical simulation task (i.e., laparoscopic peg transfer) in medical students[18]. However it remains unknown if beneficial effects would be observed for tasks with a higher level of complexity and in more experienced operators.

The objective of the present study was to conduct a randomized controlled interventional trial using a complex simulated surgical task (i.e., carotid patch-angioplasty) to test the primary hypothesis that music attenuates an operator's psychological and physiological stress response to a simulated surgical task. We also included a direct comparison of the effect of music on inexperienced (i.e., undergraduate medical students) versus experienced (i.e., trained surgeons) operators to extend and permit comparisons with previous research identifying that preferred music improves performance and reduces mental workload in inexperienced operators during the learning phase of a new and challenging surgical task[18–20]. The inclusion of such groups enhances the ecological validity of these novel investigations, while potentially informing the design of future studies examining psychophysiological stress and its mitigation in simulated surgery.

## Methods

### Ethics
The Effect of background music on STress Responses Amongst Undergraduates and Surgeons performing Simulated Surgical tasks: A randomised cross-over interventional trial (The STRAUSS study) was a non-blinded controlled crossover trial with balanced allocation carried out at the Faculty of Health and Medical Sciences, University of Auckland between April and October 2023. The STRAUSS study was designed in accordance with the CONSORT guidelines with the crossover study extension to randomised crossover trials[21,22]. This interventional trial was registered with the World Health Organisation International Clinical Trials Registry Platform (Universal Trial Number: U1111 1288-7255, 20/02/2023), the Australia and New Zealand Clinical Trials Registry (ANZCTR reference: ACTRN12623000306617, 20/3/2023), was approved by the New Zealand Health and Disability Ethics Committee (HDEC 20/CEN/57) and conformed to the standards in the Declaration of Helsinki (2013). All participants provided explicit written informed consent prior to commencement.

### Participants
Eligible participants were undergraduate medical students (MS) with little surgical experience, and senior and experienced vascular trainee surgeons or vascular consultant surgeons (VS). MS from all year groups were included (pre-clinical, clinical, or final year students). Participants were excluded if they had a significant physical impairment that inhibited dextrous task completion, current pregnancy, a significant cardiac history, or a hearing impairment in which the auditory perception of music was not possible. Participants were asked to refrain from caffeine, cardioactive medications and strenuous exercise 12 h prior to the study. Each participant had anthropometric and demographic information gathered.

### Experiment protocol and randomisation
The experiment protocol is demonstrated in Fig. 1 as a flow diagram. Participants were in the seated position with a table in front of them. After at least 5 min at rest, baseline physiological (5 min) and psychological measures were recorded. Using a Vascular Micro-Surgery Head Simulator (More Than Simulators™, Barcelona, Spain), participants were asked to perform a patch-angioplasty anastomosis using a double armed 5–0 prolene suture and a bovine pericardial patch (Xenosure, Le Maitre) to a 1.5 cm long arterial defect in a 6 mm bio-synthetic tubegraft (Omniflow, Le Maitre) that served as a carotid artery (Fig. 2). MS participants each had a period of initial training in which they were shown a five-minute video of a patch-angioplasty anastomosis and then were trained to perform the task on the model by the researcher (A.N.). VS participants were allowed to perform the task once to familiarise themselves with the surgical model during the training phase. All participants were then instructed to perform the surgical task a total of four times, randomised (by coin toss) to music or a no-music control using a block of size two. This ensured that each participant performed the task twice under each condition, once each in the first half and second half of the protocol. Participants were instructed to perform the task as neatly and quickly as possible, and an researcher observed their work while taking notes on a clipboard. A camera was set up in the corner of the room, focused on the participants hands. After the training period and after each task was completed, there was a rest period of approximately five-minutes (Fig. 2). Participants started from a resting position, and the task was considered complete when they had trimmed the final knot. Blinding was not possible with this study design, so allocation concealment was utilised. Participants were informed of the allocation at the time of commencement of the task and were kept blinded to the randomisation method. The intervention was participant selected music, with a choice of playlist, artist, or genre radio from Spotify™. Music selection was subsequently classified by Spotify™ genre allocation, or by investigator judgement when not clear. The no-music control condition was a recorded sound clip of ambient theatre noise[23]. Following the prior methodology of Oomens et al.[18], in order to minimise observer bias, both conditions were played through noise cancelling headphones at a medium volume acceptable to the participant.

### Experimental measures
The primary performance outcome measures were time to task completion, number of errors, self-rated performance on the Global Rating Scale[24] and two anastomotic quality scales: The 10-point Microsurgical Anastomosis Rating Scale (MARS10)[25] the End-Product Rating Scale (EPRS)[26].

Primary psychological outcome measures were the Six-Item State Trait Anxiety Inventory (STAI-6) and a Surgical Taskload Index (SURG-TLX) completed at baseline and immediately on task completion after each of the four trials. The STAI-6 has reported strong internal reliability, to be highly correlated with the full STAI, and
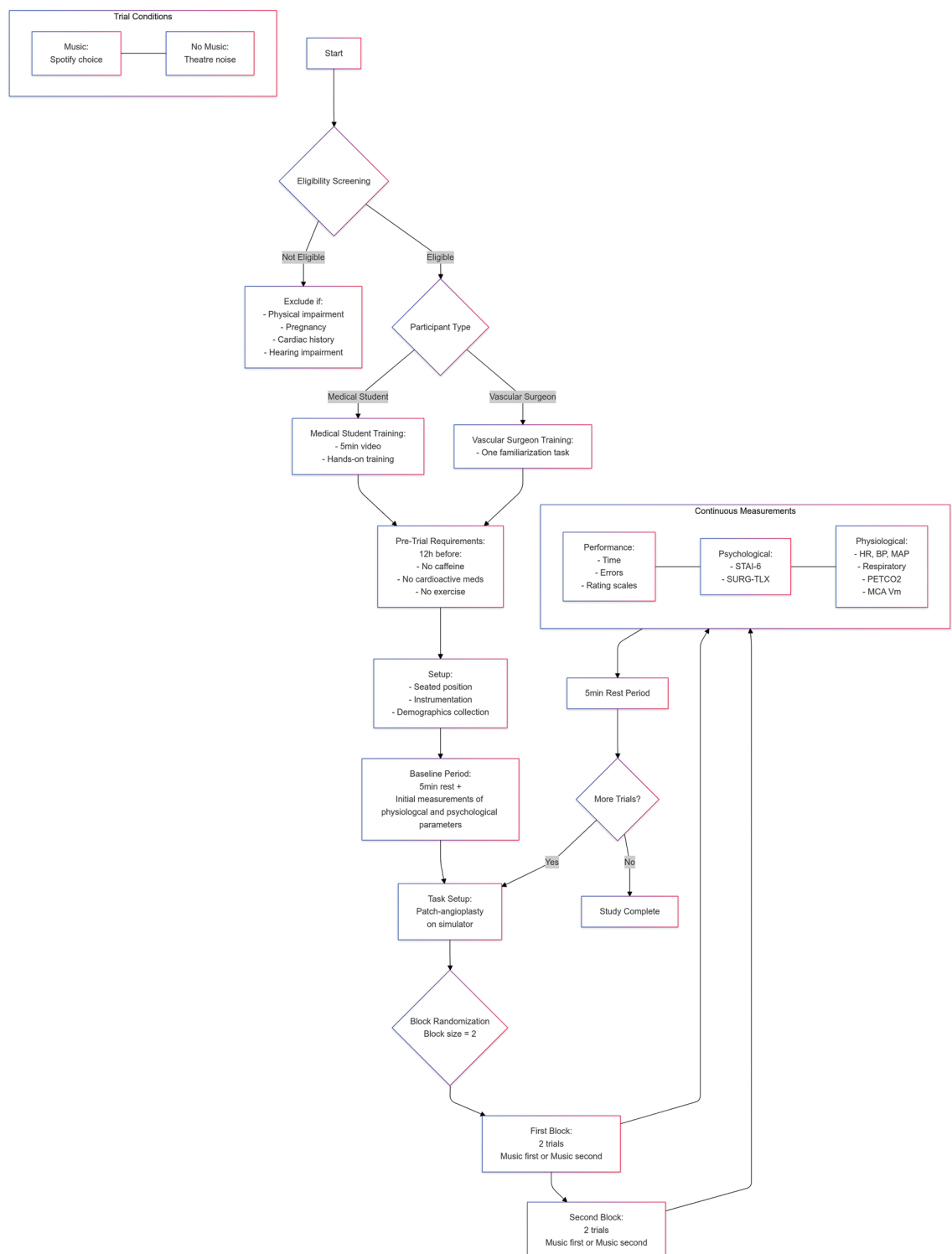
**Fig. 1.** Experiment Protocol (Participant Flow). Psychological outcome measures were the Six-Item State Trait Anxiety Inventory (STAI-6) and a Surgical Taskload Index (SURG-TLX). Physiological outcome measures were heart rate (HR) blood pressure (BP), mean arterial pressure (MAP), respiratory frequency, partial pressure of end-tidal carbon dioxide (PETCO2) and middle cerebral artery mean blood velocity (MCA Vm). Music (M), Non music (NM).

it has been used in many healthcare settings[27–29]. Each component of the STAI-6 could possibly score between 1 and 4, with a minimum total score of 6 and maximum score of 24[30]. The SURG-TLX is a multidimensional, surgery-specific workload measure that was developed and validated specifically for use in surgery from the NASA Task-load Index[31]. The highest possible score for the total SURG-TLX is 120, and the minimum score is

**Fig. 2**. Experimental set-up. Panel (**A**), the monitoring of blood pressure (BP; automated sphygmomanometer), middle cerebral artery flow velocity (MCA Vm; transcranial Doppler ultrasound), heart rate (HR; electrocardiography), respiratory frequency and partial pressure of end-tidal carbon dioxide (R$f$ and $P_{ET}CO_2$; nasal cannula). Panel (**B**), the surgical model with (from left top) suture scissors, suture holder, fine Potts scissors, Castro needle holder, Gerald forceps, head, and neck model with self-retainer within wound, green drapes to cover the handles, and (from bottom left) 5–0 prolene suture packets, artery clip forceps, basic forceps. Panel (**C**), an example of the completed carotid patch-angioplasty procedure using a 5–0 prolene suture to join a bovine pericardial patch (Xenosure) to the model artery (Omniflow).

0, with each of 6 domains scoring on a Likert scale between between 0 and 20[31]. A non-weighted raw score was generated with a higher score representing higher task-load.

Participants were instrumented for continuous monitoring of heart rate (HR) by a lead II electrocardiogram (BioAmp, FE231, ADInstruments, Bella Vista, NSW, Australia) and intermittent monitoring of blood pressure by an automated digital sphygmomanometer fixed to the left upper arm (HEM-7121, Omron Healthcare, Kyoto Japan) (Fig. 2). Mean arterial pressure (MAP) was calculated post-hoc. Respiratory frequency (R$f$) and partial pressure of end-tidal carbon dioxide ($P_{ET}CO_2$) were monitored continuously via a single-nostril nasal cannula using a gas analyzer (Respiratory Gas Analyzer, ML206, ADInstruments) calibrated using standard gases. Middle cerebral artery mean blood velocity (MCA Vm) was measured by transcranial Doppler ultrasonography (ST3, Spencer Technologies, Redmond, WA, USA). Blood velocity in the M1 segment of the middle cerebral artery was measured using a 2 MHz probe fixed over the temporal window with an adjustable headband. The M1 segment was identified using search techniques described elsewhere[32,52].

### Data analysis

Raw signals underwent analog-to-digital conversion at 1 kHz (Powerlab and LabChart v8; ADInstruments) and were stored for offline analysis. From the ECG, HR was calculated on a beat-to-beat basis. Breath-by-breath R$f$ and $P_{ET}CO_2$ were obtained, and artefacts were removed. Raw physiological responses were averaged over a 1-min period at four timepoints: rest (*3 min prior to task commencement*), the start of the task (*immediately after commencement*), the isometric midpoint of the task, and the end of the task (*immediately prior to completion*). R-R intervals were extracted from ECG data over a 5-min period at the overall resting baseline and the isometric midpoint of the task and subsequently analysed in line with the European Society of Cardiology (ESC) and the North America Society of Pacing and Electrophysiology standards of measurement of heart rate variability[33]. The ECG data were independently assessed and verified to exclude any artefacts, pauses, ectopics, and analysis was performed using HRV software (Kubios HRV Standard v3.5.0, Finland) with medium beat correction[34]. Time domain: RMSSD (square root of the mean of the sum of the squares of differences between adjacent NN intervals), frequency domain: LF:HF ratio (HF, high frequency [0.15–0.4 Hz]; LF, low frequency [0.04–0.15 Hz]) and composite scores of the parasympathetic nervous system (PNS index), sympathetic nervous system (SNS index) and Baevsky's stress indices were reported[34,35]. Poincare plot (non-linear) analysis is reported as standard deviation of the cloud of points in the direction traverse to the line-of-identity (SD1, 'short-term'), and the standard deviation of the cloud of points in the direction of the line-of-identity (SD2, 'long-term')[36].

## Statistical analysis

Investigators were blinded to the allocation of each trial during data extraction. Based on our previous work showing that mean operative SURG-TLX amongst surgeons was $48 \pm 22$, it was estimated that a total of 20 participants would be sufficient to detect a 30% difference between music and control conditions[37]. Performance outcomes were summarised and compared between trial arms with a two-tailed paired t-test. Psychological and HRV outcomes were compared with a two-way repeated measures ANOVA with group (MS, VS) and condition (baseline, music, control) as the main effects. Physiological outcomes were compared using three-way repeated measures ANOVA with group (MS, VS), condition (music, control), and time (baseline, start, midpoint, end) as main effects. Normality assumptions were checked by visually assessing residual QQ plots, which showed points on or near the reference line without systematic deviations that required model adjustments or transformations. Tukey's test and the Geisser-Greenhouse epsilon were used for post-hoc analysis of significant main effects and interactions. Statistical significance was accepted at $p < 0.05$. Pairwise deletion was used for missing data. Data are presented as mean ± standard deviation unless otherwise stated. Statistical analyses were performed using Prism 9 (Graphpad Software Inc., San Diego, California, USA)[38] and RStudio (R Core Team 2019)[39]. Most data sets were analyzed using a linear mixed-effects model conducted using the R package lmerTest[40] via RStudio.

## Results

### Protocol adherence

The protocol was well tolerated. All MS participants were able to learn and perform the surgical tasks, and all participants completed the protocol. The music intervention and control conditions did not need to be turned off for any participant.

### Participant characteristics

As shown in Table 1, the MS group (n = 15) were aged $23 \pm 2.7$ years (mean ± SD), with a height and weight of $170 \pm 10$ cm and $67 \pm 21$ kg, respectively. Ten (67%) were female and five (33%) were students who were in the first three university-based years of the undergraduate degree, and the remaining were in their final three hospital-based years. The pop genre was the most popular choice of the music intervention (n = 4), although choices also included classical, Rhythm and blues, atmospheric instrumental, jazz and metal. The VS group (n = 12) were older ($42 \pm 10$ years; $p < 0.0001$ vs. MS), taller ($179 \pm 8$ cm; $p = 0.01$), heavier ($81 \pm 12$ kg; $p = 0.04$) and more likely to be male ($p < 0.01$). There was 1 female participant and 8 were consultant surgeons. The most popular choices in the VS group were pop, easy listening and 80's music equally.

### Surgical performance

Figure 3 illustrates the mean performance measures for the surgical task. Significant group effects were observed for time to completion ($F_{(1,25)} = 43.4$, $p < 0.0001$), errors ($F_{(1,25)} = 28.9$, $p < 0.0001$), and global rating scale ($F_{(1,25)} = 16.4$, $p = 0.0004$). The VS cohort (mean difference of main effect, 95% CI) were faster to complete the task ($-10$ min [$-13.2$ to $-6.8$]), made fewer errors ($-4$ [$-7$ to $-3$]) and had higher self-rated scores (6 [2–9]). Music did not affect time to completion, number of errors or global rating scale (i.e., no condition effect was observed).

| | Medical students (MS) | Vascular surgeons (VS) |
|---|---|---|
| Demographics | | |
| Number of participants | 15 | 12 |
| Age (years)*** | 23 (18–29) | 43 (31–61) |
| Female Gender** | 10 (67%) | 1 (8%) |
| Height (cm)* | 170 (155–188) | 179 (168–196) |
| Weight (kg)* | 67 (43–116) | 81 (61–101) |
| Pre-clinical | 5 (33%) | |
| Trainees | | 4 (33%) |
| Choice of music | Pop was the most popular choice of music for the protocol (4), though the choices also included classical, R&B, atmospheric instrumental, jazz and metal | Pop, easy listening and 80's music were the most popular choices (3), others included jazz/funk, electronica and Spotify's "Good Energy" playlist |

**Table 1**. Participant characteristics. Participants included medical students (MS, n = 15) and vascular surgeons (VS, n = 12). Values are presented as mean (range) for age, height, and weight. Gender, training stage, and choice of music are expressed as counts (percentages). MS includes pre-clinical and clinical students; VS includes consultant surgeons and trainees. Music choice was self-selected for participation in the protocol. Group differences shown as *p < 0.05, **p < 0.01, ***p < 0.001.
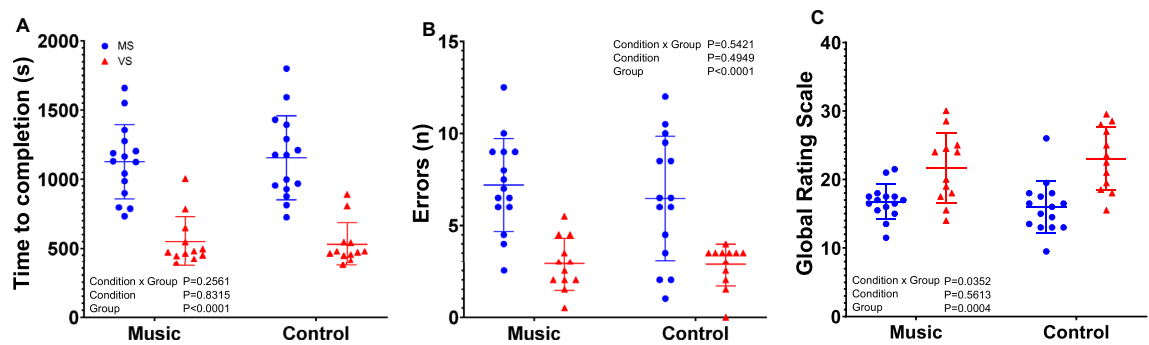
**Fig. 3**. Simulated surgical task performance measures in the music and control (non-music) conditions for the medical student (MS) and vascular surgeon (VS) groups. Symbols represent individual values for MS (blue, n = 15) and VS (red, n = 12) with horizontal bars representing means with standard deviations. *P* values show results of two-way repeated measures analysis of variance.
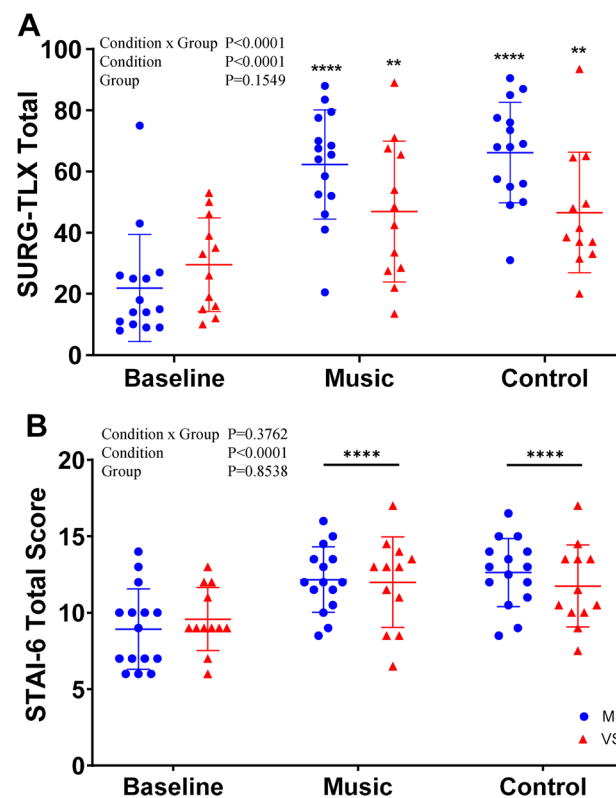


**Fig. 4**. Psychological measures at baseline and during the simulated surgical task under music and control (non-music) conditions in the medical student (MS) and vascular surgeon (VS) groups. SURG-TLX Total, Surgical Taskload Index; STAI-6, Total Six-Item State-Trait Anxiety Inventory Score. Symbols represent individual values for MS (blue, n = 15) and VS (red, n = 12) with horizontal bars representing means with standard deviations. *P* values show results of two-way repeated measures analysis of variance. Results of post-hoc analysis with Tukey HSD shown as ** = $p < 0.01$ and **** = $p < 0.0001$ versus Baseline.

### Psychological responses

SURG-TLX total score and STAI-6 total score for baseline, music, and control (condition main effect) for MS and VS (group main effect) are shown in Fig. 4 and Table 2. There was an interaction between condition and group for SURG-TLX total score (F(2,50) = 12.36, $p < 0.0001$). MS exhibited pronounced increases from baseline by Δ40 [26–55, $p < 0.0001$] with music, and by Δ44 [31–57, $p < 0.0001$] with control, that tended to be greater than VS who exhibited increases from baseline by Δ17 [7–28, $p < 0.01$] in music and by Δ17 [8–27, $p < 0.01$] in control. SURG-TLX total score was not different between music and control in either group (MS $p = 0.392$ VS $p = 1.000$). A significant main effect of condition was noted for STAI-6 total scores (F(1.9,48.3) = 18.7, $P < 0.0001$), with an

| | Baseline | | Music | | Control | | ANOVA P values | | |
|---|---|---|---|---|---|---|---|---|---|
| | MS | VS | MS | VS | MS | VS | Condition | Group | Interaction |
| STAI-6 | | | | | | | | | |
| Total score [###,††††] | 9 ± 3 | 10 ± 2 | 12 ± 3 | 12 ± 3 | 13 ± 3 | 12 ± 3 | **< 0.0001** | 0.854 | 0.376 |
| Calm [#### ††††] | 2 ± 1 | 2 ± 1 | 3 ± 1 | 2 ± 1 | 3 ± 1 | 2 ± 1 | **0.0003** | 0.837 | 0.268 |
| Relaxed [## ††††] | 2 ± 1 | 2 ± 1 | 2 ± 1 | 2 ± 1 | 3 ± 1 | 3 ± 1 | **0.0001** | 0.864 | 0.135 |
| Content [### ††] | 2 ± 1 | 2 ± 1 | 2 ± 1 | 3 ± 1 | 2 ± 1 | 3 ± 1 | **0.0003** | 0.489 | 0.789 |
| Tense [## †††] | 1 ± 1 | 2 ± 1 | 2 ± 1 | 2 ± 1 | 2 ± 1 | 2 ± 1 | **0.0002** | 0.529 | 0.967 |
| Upset [## ††] | 1 ± 0 | 1 ± 0 | 1 ± 1 | 1 ± 1 | 1 ± 1 | 1 ± 0 | **0.002** | 0.922 | 0.807 |
| Worried | 1 ± 1 | 1 ± 0 | 1 ± 1 | 1 ± 0 | 1 ± 1 | 1 ± 0 | 0.691 | 0.453 | 0.174 |
| SURG-TLX | | | | | | | | | |
| Total score [#### ††††] | 16 ± 17 | 24 ± 14 | 56 ± 17 | 14 ± 22 | 60 ± 16 | 41 ± 18 | **< 0.0001** | 0.155 | **< 0.0001** |
| Mental demands [## ††] | 3 ± 3 | 7 ± 5 | 11 ± 5 | 8 ± 5 | 11 ± 5 | 8 ± 4 | **0.002** | **0.002** | **0.017** |
| Physical demands [### ††††] | 2 ± 4 | 3 ± 2 | 8 ± 6 | 7 ± 5 | 9 ± 6 | 6 ± 4 | **< 0.0001** | **0.014** | **0.044** |
| Temporal demands [#### ††††] | 2 ± 3 | 4 ± 4 | 10 ± 4 | 9 ± 4 | 10 ± 4 | 8 ± 3 | **< 0.0001** | 0.966 | 0.414 |
| Task complexity [#### ††††] | 3 ± 5 | 3 ± 4 | 11 ± 5 | 6 ± 3 | 10 ± 4 | 6 ± 3 | **< 0.0001** | **0.016** | **0.0002** |
| Situational stress [### ††] | 3 ± 4 | 4 ± 3 | 10 ± 5 | 7 ± 4 | 11 ± 5 | 7 ± 4 | **0.0004** | **0.031** | 0.101 |
| Distractedness [## ††††] | 2 ± 3 | 3 ± 5 | 6 ± 5 | 4 ± 4 | 9 ± 6 | 7 ± 4 | **< 0.0001** | **0.006** | **0.010** |

**Table 2.** Psychological parameters at baseline and during the simulated surgical task under music and control (non-music) conditions in the medical student (MS) and vascular surgeon (VS) groups. Values are mean with standard deviations for MS (n = 15) and VS (n = 12). P values show the results of two-way repeated measures analysis of variance. Results of post-hoc analysis with Tukey HSD shown as * = $p < 0.05$ Music vs Control. # = $p < 0.05$, ## = $p < 0.01$, ### $p < 0.001$, #### = $p < 0.0001$ Baseline vs. Music. † = $p < 0.05$, †† = $p < 0.01$, ††† $p < 0.001$, †††† = $p < 0.0001$ Baseline vs Control. SURG-TLX, Surgical Taskload Index; STAI-6, Six-Item State-Trait Anxiety Inventory. Significant values are in bold.

increase from baseline to control (Δ3 [1.7–4.3], $p < 0.0001$) and baseline to music (Δ3 [1.4–4.3], $p = 0.0001$), but no difference observed between music and control (−0.15 [−1.4–1.1], $p = 0.955$).

### Physiological responses
Figure 5 demonstrates the HR, MCA Vm, $P_{ET}CO_2$, and RR responses to the surgical task at four timepoints (rest, start, midpoint and end; time main effect), in MS and VS (group main effect), for music and no music (condition main effect). HR increased from baseline at all timepoints (time main effect; $F_{(3,75)} = 14.2$, $p < 0.0001$) and reached a peak at the end of the task (Δ5.1 bpm [3.0–7.1] vs. baseline, $p < 0.0001$). HR was lower in VS (group main effect; $F_{(1,25)} = 14.8$, $p = 0.0007$) but was not affected by music (condition main effect; $p = 0.204$). Similarly, MCA Vm was increased from baseline at all timepoints during the task (time main effect; $F_{(3,75)} = 36.8$, $p < 0.0001$) and was lower in VS (group main effect; $F_{(1,25)} = 9.2$, $p = 0.006$), but was not affected by music (condition main effect; $F_{(1,25)} = 3.0$, $p = 0.094$). $P_{ET}CO_2$ increased during the task (time main effect; $F_{(3,75)} = 15.8$, $p < 0.0001$) reaching a peak at the midpoint (+ 1.5 [0.95–2.1] mmHg, $p < 0.0001$), and overall was lower with music (condition main effect; $F_{(1,25)} = 5.0$, $p = 0.034$). RR increased only at the end of the task (time main effect; $F_{(3,75)} = 2.9$, $p = 0.040$; + 1.7 [0.3–3.0] bpm vs baseline, $p < 0.01$). MAP was higher in VS than MS (group main effect; $F_{(1,25)} = 6.0$, $p = 0.022$), although not different between conditions (condition main effect; $F_{(1,25)} = 0.05$, $p = 0.822$).

### HRV responses
HRV parameters for the baseline, music, and control conditions, in the MS and VS groups, are shown in Table 3 and Fig. 6. For PNS Index, there was a main effect of condition ($F_{(1.1,27.8)} = 17.5$, $p = 0.0002$). with PNS Index being lower than baseline in both control (−0.45 [−0.73 to −0.18], $p < 0.0009$) and music (−0.57 [−0.91 to −0.23], $p < 0.0009$), and lower in music than control (−0.11 [−0.22 to −0.008], $p = 0.032$). There was also a main effect of group ($F_{(1,25)} = 5.9$, $p = 0.023$), such that MS had a lower PNS index (−0.8 [−1.5 to −0.12]) than VS. There was a main effect of condition ($F_{(1.1,28.3)} = 12.6$, $p = 0.001$) for SNS Index with values being higher than baseline during both control (0.54 [0.96–0.13], $p < 0.009$) and music (0.68 [0.22–1.14], $p < 0.003$), and higher during music than control (0.14 [0.004–0.27], $p = 0.042$). SNS index was higher in MS than VS (group main effect $F_{(1,25)} = 6.2$, $p = 0.019$; Δ + 1.0 [0.17–1.8]).

There was a main effect of condition for RMSSD ($F_{(1.2, 30.2)} = 5.6$, $p = 0.020$), with a tendency for a decrease from baseline during music (− 5.98 [− 12.4–0.48], $p = 0.073$) and control (− 3.72 [− 8.9–1.4], $p = 0.192$), and lower values during music than control (− 2.3 [− 4.7–0.24], $p = 0.081$). Total power also exhibited a main effect of condition ($F_{(1.3,32.3)} = 3.9$, $p = 0.046$), with lower values in music than control (− 322 [− 624 to − 20], $p = 0.035$). There was a main effect of condition for HF parameters ($F_{(1.3,32.2)} = 4.2$, $p = 0.038$), though no significant results on post-hoc analysis. There were no significant main effects or interactions noted for LF power and LF:HF ratio, while there was an interaction between condition and group for Baevsky's stress index ($F_{(2,50)} = 3.6$, $p = 0.035$), although no significant post hoc differences were observed.
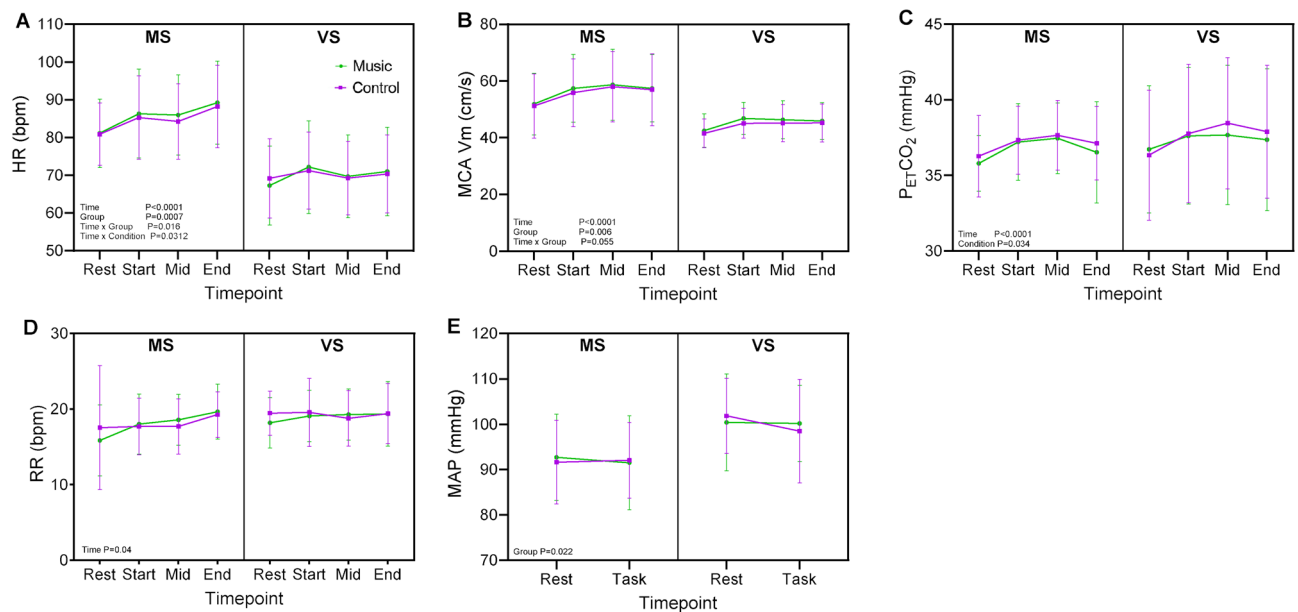
**Fig. 5**. Physiological responses to the simulated surgical task under music and control (non-music) conditions in the medical student (MS) and vascular surgeon (VS) groups. Panels show (**A**) heart rate (HR); (**B**) middle cerebral artery flow velocity (MCA Vm); (**C**) partial pressure of end-tidal carbon dioxide ($P_{ET}CO_2$, mmHg); (**D**) respiratory rate (RR); and (**E**) mean arterial pressure (MAP, mmHg). Values are expressed as mean with standard deviations for MS (green, n = 15) and VS (n = 12, purple). $P$ values show the results of three-way repeated measures analysis of variance. Results of post-hoc analysis with Tukey HSD shown as * = $p < 0.05$.

|  | Baseline | | Music | | Control | | ANOVA P values | | |
|---|---|---|---|---|---|---|---|---|---|
|  | MS | VS | MS | VS | MS | VS | Condition | Group | Interaction |
| RMSSD (ms) | 38.4 ± 16.3 | 43.4 ± 27.3 | 36.3 ± 12.7 | 32.6 ± 15.5 | 37.5 ± 12.9 | 36.2 ± 20.4 | 0.020 | 0.997 | 0.080 |
| Total power (ms²)* | 2144 ± 1427 | 2521 ± 2008 | 2012 ± 1107 | 1356 ± 853 | 2249 ± 1340 | 1783 ± 1313 | **0.045** | 0.596 | 0.069 |
| High frequency (ms²) | 788 ± 708 | 752 ± 1027 | 549 ± 394 | 434 ± 397 | 541 ± 319 | 609 ± 679 | **0.038** | 0.898 | 0.649 |
| Low frequency (ms²) | 1253 ± 835 | 1598 ± 1349 | 1262 ± 677 | 845 ± 508 | 1497 ± 809 | 1062 ± 589 | 0.139 | 0.501 | 0.055 |
| LF:HF ratio | 2.1 ± 1.5 | 3.6 ± 3.8 | 2.9 ± 1.7 | 2.8 ± 1.8 | 2.9 ± 1.7 | 3.1 ± 2.5 | 0.834 | 0.400 | 0.296 |
| Baevsky's Index | 11.1 ± 5.0 | 9.2 ± 2.8 | 10.8 ± 2.1 | 12.1 ± 4.4 | 10.7 ± 1.9 | 11.4 ± 4.3 | 0.150 | 0.971 | **0.035** |
| SDNN (ms) | 45 ± 13 | 48 ± 17 | 45 ± 12 | 36 ± 11 | 46 ± 11 | 40 ± 14 | 0.202 | 0.401 | 0.096 |
| SD1 (ms) | 27 ± 11 | 31 ± 19 | 26 ± 9 | 23 ± 11 | 27 ± 9 | 26 ± 14 | 0.258 | 0.989 | 0.503 |
| SD2 (ms) | 57 ± 15 | 60 ± 18 | 59 ± 15 | 45 ± 13 | 59 ± 13 | 49 ± 16 | 0.086 | 0.141 | **0.013** |

**Table 3**. Heart rate variability parameters at baseline and during the simulated surgical task under music and control (non-music) conditions in the medical student (MS) and vascular surgeon (VS) groups. Values are mean with standard deviations for MS (n = 15) and VS (n = 12). $P$ values show the results of two-way repeated measures analysis of variance. Results of post-hoc analysis with Tukey HSD shown as * = $p < 0.05$ Music vs Control, RMSSD, root mean square of successive differences; LF: HF ratio, Low frequency: high frequency ratio. Significant values are in bold.

## Discussion

This study examined the effect of background music on task performance, and the psychological and physiological responses to a complex simulated surgical task under stressful circumstances. The five major novel findings were, (1) music did not affect either the speed or accuracy of the simulated surgical stress task performance, (2) the task increased subjective ratings of anxiety and task load but these were unaffected by music, (3) HR, MCA Vm, $P_{ET}CO_2$, RR, and SNS index were increased, while PNS index decreased during the task, indicative of a physiological stress response, (4) PNS index was reduced and SNS index was increased in the presence of music suggestive of an arousal response, and finally (5) the more experienced group (VS) performed faster, with fewer errors and greater self-rating than the inexperienced (MS) group, and had overall lower HR, MCA Vm and MAP, but there were no psychological or physiological differences between VS and MS in their responses to music.
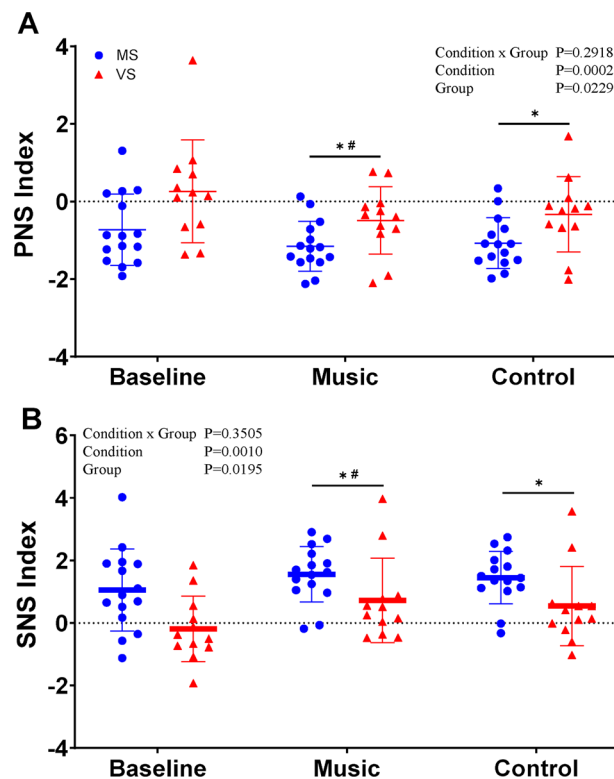
**Fig. 6.** Heart rate variability parameters at baseline and during the simulated surgical task under music and control (non-music) conditions in the medical student (MS) and vascular surgeon (VS) groups. Panels show (**A**), parasympathetic nervous system index (PNS Index) and (**B**) sympathetic nervous system index (SNS Index). Symbols represent individual values and horizontal bars show mean with standard deviations for MS (blue, n = 15) and VS (red, n = 12). *P* values show the results of two-way repeated measures analysis of variance. Results of post-hoc analysis with Tukey HSD shown as * = $p < 0.05$ Baseline, # = $p < 0.05$ versus Control.

## Task performance

A beneficial effect of music on simulated surgical task performance was not demonstrated in the present study in either inexperienced or experienced operators. This is out of keeping with El Boghdady's systematic review that concluded classical music at a low to medium volume improves accuracy and speed, and outweighs negative effects of distractedness, on the basis of n = 18 studies that were predominantly simulation studies and surveys[41]. This may in part be due to the novel use of a complex, non-laparoscopic surgical model of carotid patch-plasty in this field of research, a task that required the performance of multiple surgical steps (i.e. knot tying, suturing, patch trimming) which may take variable lengths of time. To investigate this variable, future research may consider utilising a study design that includes a complex multi-step task in comparison to a simple task. Music has also been shown to hasten the learning process of surgical tasks on the Da Vinci robot[19]. In the present study, there was no music or sound during the training phase of the protocol, and future research could investigate the effect of music on the learning trajectory of complex surgical tasks and speed of skill acquisition.

## Psychological stress

Despite the simulated carotid patch-angioplasty task inducing a clear psychological task load and anxiety response from operators, this was not modified by music in either the VS or MS group. The unweighted version of the SURG-TLX was chosen as the added time of the longer weighted questionnaire would have significantly increased the length of time the protocol would take, given its repeated nature. While previous studies have shown high correlation between the weighted and unweighted scores[42], there is a potential reduction in the generalisability or comparability of our findings to similar studies that have utilised the weighted version of the SURG-TLX. Elevated SURG-TLX scores may, in the real world setting, be associated with longer length on bypass in cardiac surgery, and with particular surgical team roles[43,44].

Background music decreased workload scores in two previous studies of inexperienced operators performing simulated laparoscopic surgery[17,18]. The possible reasons for the conflicting findings of the present study include the nature of the control condition (silence vs. ambient OR noise) and the nature of simulated task. SURG-TLX measured reproducible changes from baseline in the participants in this study. Other authors have preliminary recommended the SURG-TLX as a standard measure to assess operators' experience in studies of surgical innovation, as it has been found to be most relevant, comprehensive, and comprehensible[45]. STAI-6 scores increased during task performance (around 3 points in both groups), though the level of experience did not affect this increase. STAI-6 has been frequently used validated scale for stress measurement and has successfully

been demonstrated to correlate to LF:HF though a direct comparison between STAI-6 and SURG-TLX has yet to be published[46].

## Physiological stress

The data here represents the most detailed physiological phenotype of a surgeon performing a simulated task to date[10,47]. In other studies, during surgery, mean HR has been measured to be the most consistent biomarker elevated by 6 to 22 bpm from baseline compared to 5 bpm in this study[47]. In our study, there was no difference in HR between music and control conditions. Analysis of HRV (i.e., SDNN, RMSSD, LF, HF, HF:LF ratio) has been shown to be a good objective assessment of mental stress in the surgical setting[48]. Notably, we observed that music reduced the PNS index and increased the SNS index, suggestive of an arousal response, which is contrary to widespread perception of OR music, and previous interventional studies[7,8,10,18]. Further evidence for an arousal response to music is the lowering of $P_{ET}CO_2$, which would fit with an increase in minute ventilation under the music condition. Allen et al. suggested that self-selected music could reduce skin conductance, pulse and blood pressure and improve speed and accuracy of arithmetic tasks when compared to control music or the absence of music[49]. One study reported an increase in MAP between control and self-selected music[50], though four other studies also demonstrated no change in HR or MAP under music conditions compared to control[10]. Finally, relaxing music (slow-paced classical pieces) when compared to control showed greater return to baseline in high frequency HRV after a stressing event suggesting a role for music in the recovery for surgeons after stressful operations[51].

Transcranial Doppler ultrasound was first used by Aaslid et al[52] to measure cerebral blood velocity in dynamic cerebral autoregulation research, and its application to surgeons is novel. Cerebral perfusion (i.e., MCA Vm) was significantly increased during task performance, with maximal increase during the midpoint of the operation, although a difference between music and control conditions was not observed. The cerebrovascular changes during task performance reported in this study are not surprising, in that MCA Vm has been shown to increase bilaterally in mathematical cognitive effort[53], and during simulated robot assisted surgery, the pre-frontal cortex (supplied by the MCA and anterior cerebral artery) differentially activates[54,55]. The functional hyperaemia we observed may be linked to neuronal activation during the surgical task (i.e., neurovascular coupling), but the coincident elevation in $P_{ET}CO_2$ (a proxy for arterial $CO_2$ concentration a powerful cerebrovascular vasodilator) during task performance is also likely to be a contributory factor[56]. The nature of the task meant that non-invasive beat-to-beat blood pressure could not be measured using finger photoplethysmography, hence cerebral perfusion pressure estimates cannot be incorporated into the interpretation of the MCA Vm response[57]. It should be noted that transcranial Doppler assesses artery blood velocity, rather than blood flow, and whether the this is representative of cerebral blood flow rests on the assumption that vessel diameter remains constant.

## The impact of experience

The VS cohort were older, taller, heavier, and contained more males which may account for the baseline group main effects between groups (i.e. lower HR, lower MCA Vm, higher MAP). Unsurprisingly, the greater experience was associated with improved speed and quality of the task performance[58,59]. A time x group interaction in HR is suggestive of attenuated cardiovascular reactivity to the task in the VS group, for whom the surgical task would have been more familiar. This is in line with previous work that has shown senior surgeons to display a better stress management capacity and are less prone to cognitive distraction compared to junior surgeons[60]. Future work should then consider that if music were to attenuate a psychophysiological response to surgical tasks, this effect may be less detectable in an experienced population where the stress response to tasks is smaller.

## The type of music

The question often arises as to whether the type of music is important. Self-selected music is a commonly researched music choice in the surgical setting, in part due to Allen et al.'s seminal paper, and was therefore used in this study[10,49]. Music considered to be unpleasant may generate poorer surgical performance, an effect which may be measurable through application of electro-encephalography[61]. Dissonance in music has been shown to induce unpleasant feelings and physiological responses which may lend support to the application of consonant music[62]. Furthermore, faster music may reduce OR preparation time[63], increase the rate of task performance (while also increasing the number of mistakes)[64], greater sound intensity can increase perceived arousal and high rhythmic tension may more strongly affect HRV[65], which was chosen (rather than silence) for ecological validity.

## Strengths and limitations

Several experimental factors should be considered when interpreting the results of the present investigation. The crossover design was counterbalanced to minimise possible learning effects and all participants undertook a supervised training and practice period prior to being assessed to further mitigate against learning effects. To increase the experience of stress during the surgical task several experimental manipulations were employed. The Trier Social Stress Test is an ecologically valid stressor induced by public speaking, and in the present study and social evaluation was incorporated with the presence of an unresponsive yet engaged assessor and a camera setup (without filming)[66]. Perceived time pressure refers to the sense of time pressure experienced in the absence of genuine time limits and has been shown to be an empirical stressor, and was also incorporated into the study design[67]. Finally, participants were asked to complete task to the highest quality possible to maintain a level of sustained mental effort investment during the protocol in an aim to increase mental workload, stress and fatigue during the task[68]. The results indicate that the task and the applied manipulation were successful in generating a stress response. It is rarely silent in a working operating room. Previous interventional studies are heterogenous in the choice of control condition, however in this study ambient theatre noise was selected as the control

condition, rather than silence for ecological validity[10,69]. As such it is possible that previous work showing that music improved simulated laparoscopy compared to a no music / silence control may be due to the presence of auditory stimulation (or the absence of silence), rather than the music itself[18]. HRV analyses were used to provide an indirect estimate of cardiac autonomic behaviour. However, there remains controversy surrounding the physiological interpretation of these metrics[70]. Furthermore, we acknowledge that the composite HRV scores reported (i.e., SNS and PNS indices) are currently unconventional[33,34] but believe that they are helpful in summarising related metrics and we have also included the component metrics in Table 3.

## Conclusion

We observed that the performance of a simulated carotid patch-angioplasty under stressing conditions successfully increased markers of psychological and physiological stress in both inexperienced and experienced operators. Despite previous research identifying generally positive surgeon perceptions, background music did not improve either surgical task performance or attenuate subjective ratings of task load and anxiety, but in fact evoked additional physiological arousal evidenced by decreased PNS index and increased SNS index. As expected, the more experienced group performed faster and more accurately than the inexperienced group, but there were no psychological or physiological differences in their responses to music.

## Data availability

The datasets generated and analysed during the current study are not publicly available due participant confidentiality but de-identified and anonymous data are available from the corresponding author upon reasonable request.

## References

1. Arora, S., Flindall, I. & Youngson, G. G. Performance-shaping factors. In *Enhancing surgical performance: A primer in non-technical skills* (eds Flin, R. et al.) (CRC Press, 2015).
2. Arora, S. et al. Factors compromising safety in surgery: stressful events in the operating room. *Am. J. Surg.* **199**(1), 60–65. https://doi.org/10.1016/j.amjsurg.2009.07.036 (2010).
3. Arora, S. et al. The impact of stress on surgical performance: a systematic review of the literature. *Surgery* **147**(3), 318-30-630 e1-6. https://doi.org/10.1016/j.surg.2009.10.007 (2010).
4. Sharma, A. et al. Stress and burnout in colorectal and vascular surgical consultants working in the UK National Health Service. *Psychooncology* **17**(6), 570–576. https://doi.org/10.1002/pon.1269 (2008).
5. Ullmann, Y. et al. The sounds of music in the operating room. *Injury* **39**(5), 592–597. https://doi.org/10.1016/j.injury.2006.06.021 (2008).
6. Hawksworth, C., Asbury, A. J. & Millar, K. Music in theatre: Not so harmonious. A survey of attitudes to music played in the operating theatre. *Anaesthesia* **52**(1), 79–83 (1997).
7. Narayanan, A. & Gray, A. R. First, do no harmony: An examination of attitudes to music played in operating theatres. *New Z. Med. J.* **131**(1480), 68–74 (2018).
8. Narayanan, A. et al. Broad responses and attitudes to having music in surgery (the BRAHMS study)—A South African perspective. *South Afr J Surg Suid-Afrikaanse tydskrif vir chirurgie* **61**(1), 30–38 (2023).
9. Narayanan, A. et al. Broad responses and attitudes to having music in surgery (The BRAHMS Study): An Australia and Aotearoa New Zealand Perspective. *Surg. Open Sci.* **17**, 30–34. https://doi.org/10.1016/j.sopen.2023.12.009 (2024).
10. Narayanan, A. et al. The effect of background music on stress in the operating surgeon: Scoping review. *BJS Open* **6**(5), zrac112. https://doi.org/10.1093/bjsopen/zrac112 (2022).
11. Faraj, A. A. et al. LISTEN WHILE YOU WORK? The attitude of healthcare professionals to music in the OR. *Ornac J* **33**(2), 31-2–34-50 (2015).
12. Fu, V. X. et al. The perception and attitude toward noise and music in the operation room: A systematic review. *J. Surg. Res.* **263**, 193–206 (2021).
13. Makama, J. G., Ameh, E. A. & Eguma, S. A. Music in the operating theatre: Opinions of staff and patients of a Nigerian teaching hospital. *Afr. Health Sci.* **10**(4), 386–389 (2010).
14. George, S. et al. Influence of music on operation theatre staff. *J. Anaesthesiol. Clin. Pharmacol.* **27**(3), 354–357 (2011).
15. Way, T. J. et al. Effect of noise on auditory processing in the operating room. *J. Am. Coll. Surg.* **216**(5), 933–938. https://doi.org/10.1016/j.jamcollsurg.2012.12.048 (2013).
16. Cheriyan, S. et al. The impact of operating room noise upon communication during percutaneous nephrostolithotomy. *J. Endourol.* **30**(10), 1062–1066. https://doi.org/10.1089/end.2016.0498 (2016).
17. Fu, V. X. et al. The effect of preferred music on mental workload and laparoscopic surgical performance in a simulated setting (OPTIMISE): A randomized controlled crossover study. *Surg. Endosc.* **35**, 5051–5061 (2020).
18. Oomens, P., Fu, V. X., Kleinrensink, V. E. *et al.* The effects of preferred music on laparoscopic surgical performance: A randomized crossover study. *World J. Surg* 2020:1–6.
19. Siu, K. C. et al. The effect of music on robot-assisted laparoscopic surgical performance. *Surg. Innov.* **17**(4), 306–311. https://doi.org/10.1177/1553350610381087 (2010).
20. Gil, K. et al. Satisfaction or distraction: Exposure to nonpreferred music may alter the learning curve for surgical trainees. *J. Surg. Educ.* **77**(6), 1370–1376 (2020).
21. Schulz, K. F., Altman, D. G. & Moher, D. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ* **340**, c332. https://doi.org/10.1136/bmj.c332 (2010).
22. Dwan, K. et al. CONSORT 2010 statement: extension to randomised crossover trials. *BMJ* **366**, l4378. https://doi.org/10.1136/bmj.l4378 (2019).
23. Ambience Laboratory. Surgery Background Sounds 2. ASMR Ambience Hospital Youtube2022 [Available from: https://www.youtube.com/watch?v=s4HyneDVXi0.
24. Doyle, J. D., Webber, E. M. & Sidhu, R. S. A universal global rating scale for the evaluation of technical skills in the operating room. *Am. J. Surg.* **193**(5), 551–555. https://doi.org/10.1016/j.amjsurg.2007.02.003 (2007).
25. Stogowski, P. et al. Microsurgical anastomosis rating scale (MARS10): A final product scoring system for initial microsurgical training. *Plast. Surg. (Oakv)* **29**(4), 243–249. https://doi.org/10.1177/2292550320969649 (2021).

26. Price, J. et al. A randomized evaluation of simulation training on performance of vascular anastomosis on a high-fidelity in vivo model: The role of deliberate practice. *J. Thorac. Cardiovasc. Surg.* **142**(3), 496–503. https://doi.org/10.1016/j.jtcvs.2011.05.015 (2011).

27. Marteau, T. M. & Bekker, H. The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br. J. Clin. Psychol.* **31**(3), 301–306. https://doi.org/10.1111/j.2044-8260.1992.tb00997.x (1992).

28. Court, H., Greenland, K. & Margrain, T. Measuring patient anxiety in primary care: Rasch analysis of the 6-item Spielberger State Anxiety Scale. *Value Health J Int Soc Pharmacoeconomics Outcomes Res.* **13**, 813–819. https://doi.org/10.1111/j.1524-4733.2010.00758.x (2010).

29. Bayrampour, H. et al. Reliability and validity of three shortened versions of the State Anxiety Inventory scale during the perinatal period. *J. Psychosom. Obstet. Gynaecol.* **35**(3), 101–107. https://doi.org/10.3109/0167482x.2014.950218 (2014).

30. Chlan, L., Savik, K. & Weinert, C. Development of a shortened state anxiety scale from the Spielberger State-Trait Anxiety Inventory (STAI) for patients receiving mechanical ventilatory support. *J. Nurs. Meas.* **11**(3), 283–293. https://doi.org/10.1891/jnum.11.3.283.61269 (2003).

31. Wilson, M. R. et al. Development and validation of a surgical workload measure: the surgery task load index (SURG-TLX). *World J. Surg.* **35**(9), 1961–1969. https://doi.org/10.1007/s00268-011-1141-4 (2011).

32. Willie, C. K. et al. Utility of transcranial Doppler ultrasound for the integrative assessment of cerebrovascular function. *J. Neurosci. Methods.* **196**(2), 221–237. https://doi.org/10.1016/j.jneumeth.2011.01.011 (2011).

33. European Society of Cardiology. Heart rate variability: Standards of measurement, physiological interpretation and clinical use task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation* **93**(5), 1043–1065 (1996).

34. Tarvainen, M. P. et al. Kubios HRV–heart rate variability analysis software. *Comput. Methods Programs Biomed.* **113**(1), 210–220. https://doi.org/10.1016/j.cmpb.2013.07.024 (2014).

35. Methodical Recommendations—Use Kardivar System for Determination of the Stress Level and Estimation of the Body Adaptability-Standards of Measurements and Physiological Interpretation (2009).

36. Sassi, R. et al. Advances in heart rate variability signal analysis: Joint position statement by the e-Cardiology ESC Working Group and the European Heart Rhythm Association co-endorsed by the Asia Pacific Heart Rhythm Society. *EP Europace* **17**(9), 1341–1353. https://doi.org/10.1093/europace/euv015 (2015).

37. Narayanan, A., Fisher, J. P. & Khashram, M. The MOSART Study—The effect of music on the operating surgeon—A randomised trial. In: *Australia and New Zealand Society of Vascular Surgery—Annual Scientific Conference* (ed. ANZSVS) (Gold Coast, Australia, 2021).

38. Prism G. GraphPad Prism. *GraphPad Software Inc* 2020

39. R: A Language and Environment for Statistical Computing [program]. Vienna, Austria: R Foundation for Statistical Computing (2023).

40. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: Tests in linear mixed effects models. *J. Stat. Softw.* **82**(13), 1–26 (2017).

41. El Boghdady, M. & Ewalds-Kvist, B. M. The influence of music on the surgical task performance: A systematic review. *Int. J. Surg.* **73**, 101–112. https://doi.org/10.1016/j.ijsu.2019.11.012 (2020).

42. Ma, J., McCrory, B. & Claudio, D. Comparison of weighted and unweighted task load indices for single-site surgery tasks. *Proc. Int. Symp. Hum. Factors Ergon. Health Care* **10**(1), 142–146. https://doi.org/10.1177/2327857921101080 (2021).

43. Yu, D. et al. Quantifying intraoperative workloads across the surgical team roles: Room for better balance?. *World J. Surg.* **40**(7), 1565–1574. https://doi.org/10.1007/s00268-016-3449-6 (2016).

44. Kennedy-Metz, L. R. et al. Surgery task load index in cardiac surgery: Measuring cognitive load among teams. *Surg. Innov.* **27**(6), 602–607. https://doi.org/10.1177/1553350620934931 (2020).

45. McNair, A. G. K. et al. A standardized measurement instrument was recommended for evaluating operator experience in complex healthcare interventions. *J. Clin. Epidemiol.* **153**, 55–65. https://doi.org/10.1016/j.jclinepi.2022.10.006 (2023).

46. Tjønnås, M. S. et al. Stress in surgical educational environments: A systematic review. *BMC Med Educ* **22**(1), 791. https://doi.org/10.1186/s12909-022-03841-6 (2022).

47. Budden, A. et al. A systematic review of biological changes in surgeons' acute stress levels during surgery. *Surg. Pract. Sci.* **13**, 100174. https://doi.org/10.1016/j.sipas.2023.100174 (2023).

48. The, A. F. et al. Heart rate variability as a measure of mental stress in surgery: A systematic review. *Int Arch Occup Environ Health* **93**(7), 805–821. https://doi.org/10.1007/s00420-020-01525-6 (2020).

49. Allen, K. & Blascovich, J. Effects of music on cardiovascular reactivity among surgeons. *JAMA* **272**(11), 882–884 (1994).

50. Mitta, N. et al. Surgical notes: To play or not to play. *World J. Surg.* **43**(11), 2740–2746 (2019).

51. Baccarani, A. et al. Relaxing effects of music and odors on physiological recovery after cognitive stress and unexpected absence of multisensory benefit. *Psychophysiology* **60**(7), e14251. https://doi.org/10.1111/psyp.14251 (2023).

52. Aaslid, R., Markwalder, T. M. & Nornes, H. Noninvasive transcranial Doppler ultrasound recording of flow velocity in basal cerebral arteries. *J. Neurosurg.* **57**(6), 769–774. https://doi.org/10.3171/jns.1982.57.6.0769 (1982).

53. Montoro, C. I., Duschek, S. & Reyes del Paso, G. A. Variability in cerebral blood flow velocity at rest and during mental stress in healthy individuals: Associations with cardiovascular parameters and cognitive performance. *Biol. Psychol.* **135**, 149–158. https://doi.org/10.1016/j.biopsycho.2018.04.005 (2018).

54. Hannah, T. C. et al. Neuromonitoring correlates of expertise level in surgical performers: A systematic review. *Front. Hum. Neurosci.* **16**, 705238. https://doi.org/10.3389/fnhum.2022.705238 (2022).

55. Izzetoglu, K. et al. Studying brain activation during skill acquisition via robot-assisted surgery training. *Brain Sci.* **11**(7), 937. https://doi.org/10.3390/brainsci11070937 (2021).

56. Purkayastha, S. & Sorond, F. Transcranial Doppler ultrasound: Technique and application. *Semin. Neurol.* **32**(4), 411–420. https://doi.org/10.1055/s-0032-1331812 (2012).

57. Fan, J.-L. et al. Integrative physiological assessment of cerebral hemodynamics and metabolism in acute ischemic stroke. *J. Cereb. Blood Flow Metab.* **42**(3), 454–470. https://doi.org/10.1177/0271678x211033732 (2022).

58. Rothbaum, D. L. et al. Task performance in stapedotomy: Comparison between surgeons of different experience levels. *Otolaryngol. Head Neck Surg.* **128**(1), 71–77. https://doi.org/10.1067/mhn.2003.13 (2003).

59. Loukas, C., Gazis, A. & Kanakis, M. A. Surgical performance analysis and classification based on video annotation of laparoscopic tasks. *JSLS J. Soc. Laparoendosc. Surg.* **24**(4), e2020-00057. https://doi.org/10.4293/jsls.2020.00057 (2020).

60. Marrelli, M. et al. Correlation between Surgeon's experience, surgery complexity and the alteration of stress related physiological parameters. *PLoS ONE* **9**(11), e112444. https://doi.org/10.1371/journal.pone.0112444 (2014).

61. Sammler, D. et al. Music and emotion: Electrophysiological correlates of the processing of pleasant and unpleasant music. *Psychophysiology* **44**(2), 293–304. https://doi.org/10.1111/j.1469-8986.2007.00497.x (2007).

62. Dellacherie, D. et al. The effect of musical experience on emotional self-reports and psychophysiological responses to dissonance. *Psychophysiology* **48**(3), 337–349. https://doi.org/10.1111/j.1469-8986.2010.01075.x (2011).

63. Mosaed, S., Vahidi, R. & Lin, K. Y. Effect of music tempo on operating room preparation time. *J. Perioper. Pract.* **30**(5), 141–144 (2020).

64. Dalton, B. & Behm, D. Effects of noise and music on human and task performance: A systematic review. *Occup. Ergon.* **7**, 143–152 (2007).

65. Mikutta, C. A. et al. Music, perceived arousal, and intensity: Psychophysiological reactions to Chopin's "Tristesse". *Psychophysiology* **50**(9), 909–919. https://doi.org/10.1111/psyp.12071 (2013).
66. Allen, A. P. et al. The trier social stress test: Principles and practice. *Neurobiol. Stress* **6**, 113–126. https://doi.org/10.1016/j.ynstr.2016.11.001 (2017).
67. Sussman, R. F. & Sekuler, R. Feeling rushed? Perceived time pressure impacts executive function and stress. *Acta Physiol. (Oxf)* **229**, 103702. https://doi.org/10.1016/j.actpsy.2022.103702 (2022).
68. Fairclough, S. & Mulder, B. Psychophysiological processes of mental effort investment. In *How motivation Affects Cardiovascular Response: Mechanisms and Applications* (eds Wright, R. A. & Gendolla, G. H. E.) 61–76 (American Psychological Association, 2012).
69. Katz, J. D. Noise in the operating room. *Anesthesiology* **121**(4), 894–898. https://doi.org/10.1097/aln.0000000000000319 (2014).
70. Berntson, G. G. et al. Heart rate variability: Origins, methods, and interpretive caveats. *Psychophysiology* **34**(6), 623–648. https://doi.org/10.1111/j.1469-8986.1997.tb02140.x (1997).

## Acknowledgements

## Author contributions

AN: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing—Original Draft, Writing—Review & Editing, Visualization, Supervision, Project administration, Funding acquisition MK: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing—Original Draft, Writing—Review & Editing, Visualization, Supervision, Project administration, Funding acquisition, JPF: Conceptualization, Methodology, Software, Formal analysis, Investigation, Resources, Data Curation, Writing—Original Draft, Writing—Review & Editing, Supervision, Project administration, Funding acquisition. All authors reviewed the manuscript.

## Funding

## Declarations

### Competing interests

The authors declare no competing interests.

### Ethics approval

This interventional trial was registered with the World Health Organisation International Clinical Trials Registry Platform (Universal Trial Number: U1111 1288–7255) and the Australia New Zealand Clinical Trials Registry (ANZCTR reference: ACTRN12623000306617). Ethics approval was granted by the New Zealand Health and Disability Ethics Committee (HDEC 20/CEN/57 2023 AM 9287).

### Additional information

**Correspondence** and requests for materials should be addressed to A.N.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.