

The influence of music on technical skill in robotic surgery among general surgery residents: a simulation-based experimental study

Guilherme Corvello Dias

gui.corvello@gmail.com

Faculdade de Medicina de São José do Rio Preto

Ramon Jorge Guiu

Faculdade de Medicina de São José do Rio Preto

Pedro Henrique Alves Silva

Faculdade de Medicina de São José do Rio Preto

Heitor Zancheta Andrade

Faculdade de Medicina de São José do Rio Preto

Rodrigo Tavore Strasser

Faculdade de Medicina de São José do Rio Preto

Carlos Dario da Silva Costa

Faculdade de Medicina de São José do Rio Preto

Marco Antônio Ribeiro Filho

Faculdade de Medicina de São José do Rio Preto

Gabriela Gouvea Silva

Faculdade de Medicina de São José do Rio Preto

Research Article

Keywords: Music, Robotic surgery, Simulation training, Surgical education, Medical residency

Posted Date: October 18th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11353315/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Music is commonly played in operating rooms, but its effect on the acquisition of robotic surgical skills is unclear, particularly when the music is chosen by the trainee. This study evaluated whether a personalized musical intervention influences the technical performance of general surgery residents on a robotic surgery simulator. Thirty-two residents with no previous robotic experience performed three exercises (Pick and Place, Ring and Rail 1 and Energy Dissection 1) on the RobotiX Mentor simulator. Each participant completed a first run without music and a second run with self-selected preferred music, followed by a third run under a randomly allocated condition: no music, preferred music or random music. Objective simulator metrics were summarized descriptively, and the three conditions were compared in the randomly allocated run using the Kruskal–Wallis test with Bonferroni correction for 26 comparisons. No comparison reached statistical significance after correction. The lowest unadjusted p values were observed for instrument collisions in Pick and Place ($p = 0.045$) and inappropriately applied energy in Energy Dissection 1 ($p = 0.029$; corrected $p = 0.753$). In an exploratory analysis, execution time decreased from run 1 to run 2 in all exercises ($p \leq 0.003$), consistent with learning. Participants selected 31 different songs by 25 artists. Under the conditions evaluated, personalized music was not associated with a significant improvement in technical performance during early robotic simulator training. The protocol proved feasible and reproduces the individualized way in which music is used in clinical practice; larger, fully randomized studies are warranted.

Introduction

Robotic surgery has become a major technological development in minimally invasive surgery, overcoming some of the limitations of conventional laparoscopy. Its incorporation into clinical practice has made evident the need for training methods capable of developing specific technical skills before patient care. Unlike conventional laparoscopy, robotic surgery requires adaptation to a new surgeon–platform interface, distinct visuospatial coordination and mastery of highly precise movements. In this context, robotic surgery simulators have come to play a central role in surgical training, enabling repetitive, safe and standardized practice without risk to patients [1, 2], while providing objective metrics of performance and efficiency. These measures allow training progression to be tracked and interventions that might optimize technical learning to be evaluated.

Among the environmental factors capable of influencing performance during surgical activities, music has attracted growing interest over the past decades. Its use is common in operating rooms and has been associated with an improved work environment, reduced occupational stress and increased concentration among the surgical team [3, 4]. Several studies have evaluated the impact of music on technical performance during simulated surgical tasks. Some observed improved technical efficiency, shorter execution time and lower cognitive load during laparoscopic and robotic tasks [5–9], whereas others found no significant differences between conditions with and without music [10–13].

Historically, most studies have used standardized music, such as classical music or playlists selected by the researchers. Although this strategy facilitates experimental standardization, it reduces the ecological validity of the findings, because it does not reproduce the everyday surgical environment, in which the choice of music is generally driven by the individual preferences of the team [7]. More recently, attention has turned to preferred music, on the premise that the effects of music on cognitive and motor performance depend more on familiarity, preference and the emotional meaning attributed by the individual than on the musical genre itself [7–14]. Nevertheless, studies investigating individualized musical interventions during robotic surgery training with objective simulator metrics remain scarce, particularly among general surgery residents without previous robotic experience, and the musical interventions described in the literature are poorly characterized.

The aim of this study was therefore to evaluate the influence of a personalized musical intervention on the technical performance of general surgery residents during robotic simulator training. We compared no music, the participant's preferred music and random music using the objective metrics automatically recorded by the simulator, and characterized the musical repertoire chosen by the participants.

Methods

Study design and setting

This prospective experimental study was conducted at the Surgical Simulation Laboratory of the Faculty of Medicine of São José do Rio Preto (FAMERP), Brazil, between May 2025 and February 2026, using the RobotiX Mentor virtual reality simulator (Surgical Science, Gothenburg, Sweden), which records objective performance metrics during simulated surgical tasks [15]. The study was approved by the institutional Research Ethics Committee (CAAE 86407625.8.0000.5415), and all participants provided written informed consent before data collection.

Participants

All residents enrolled in the FAMERP General Surgery Residency Program were invited, and all 32 agreed to participate, with no refusals. The only eligibility criterion was the absence of previous experience in robotic surgery; none of the participants had previous training on the simulator used in the study. No other participant characteristics were recorded.

Simulator exercises

All participants received standardized instructions on the operation of the simulator and the execution of the exercises through the same explanatory video. No warm-up or familiarization run was performed. Three exercises of increasing difficulty were used: *Pick and Place*, in which coloured pegs are transferred to the tray of the corresponding colour; *Ring and Rail 1*, in which a ring is guided along a curved metal rail

while the camera is moved simultaneously; and *Energy Dissection 1*, in which a main blood vessel is isolated from its branches using energy instruments.

Experimental protocol

Each participant performed three consecutive runs of the three exercises. The first two runs followed a fixed order: run 1 without music and run 2 with the participant's preferred music, chosen in advance. For run 3, the condition (no music, preferred music or random music) was drawn separately for each exercise using sealed envelopes, so a participant could perform different exercises under different conditions. In the random music condition, the song was drawn from a playlist previously compiled by the investigator, containing a mix of songs and musical genres; whether the participant was familiar with the song was not assessed. Music was delivered through headphones at a volume chosen by each participant and played continuously until the end of the run, restarting whenever the run lasted longer than the song. Runs were performed consecutively, without rest intervals. Location, equipment, lighting, simulator position and instructions were kept constant throughout the experiment.

Outcome measures

The outcomes were the objective metrics recorded automatically by the simulator software at the end of each exercise: execution time, economy of motion, workspace, instrument collisions, instruments out of view, grasper contact warnings, excessive instrument force and drops (pegs or rings), together with metrics specific to *Energy Dissection 1*: inappropriately applied energy, blood loss and injured vessels. Execution time, grasper contact warnings, excessive instrument force and inappropriately applied energy were recorded in seconds; economy of motion, workspace and instruments out of view in centimetres; blood loss in millilitres; and collisions, drops and injured vessels as counts. For each preferred song, the title, artist, genre, origin (Brazilian or international), presence of lyrics, perceived tempo and perceived energy were recorded. Perceived tempo and energy were reported by each participant for the song they had chosen.

Statistical analysis

Analyses were performed using R version 4.4.1 [16] in two stages. In the first stage, performance under each condition was described using the mean, standard deviation, median, minimum and maximum of each metric, separately for each exercise. These descriptive data pooled all runs performed under each condition; because condition and run order are partly confounded in this design (no music always in run 1, random music only in run 3), they were not used for inference. Mean metrics for each condition were normalized and displayed in radar charts using the *fmsb* package, version 0.7.6 [17].

In the second stage, the three conditions were compared using only the randomly allocated run (run 3), in which the condition was drawn separately for each exercise. The Kruskal–Wallis test was applied separately to each exercise–metric combination, as a non-parametric test suitable for small samples and non-normally distributed data. A total of 26 tests were performed (8 for Pick and Place, 8 for Ring and Rail 1 and 10 for Energy Dissection 1). The Bonferroni correction was applied by multiplying each p value by the number of tests, and differences were considered significant when the corrected p value was below 0.05. Effect sizes were estimated as epsilon squared ($\epsilon^2 = H/(n - 1)$). Distributions of selected metrics by condition were displayed in box plots using the ggplot2 package, version 4.0.3 [18]. As an exploratory analysis, execution time in run 1 (no music) and run 2 (preferred music) was compared within participants for each exercise using the Wilcoxon signed-rank test; because these runs differed in both order and musical condition, this comparison reflects the combined effect of repetition and music and was not used to test the effect of music.

Results

Participants and musical intervention

All 32 invited residents completed the protocol. There were no missing data. In run 3, the 96 attempts comprised 27 without music, 33 with preferred music and 36 with random music (Pick and Place, 10/10/12; Ring and Rail 1, 9/12/11; Energy Dissection 1, 8/11/13, for no music/preferred/random music, respectively). The 32 participants chose 31 different songs (one song was chosen by two participants) performed by 25 artists. Rock/alternative (9; 28.1%), sertanejo (8; 25.0%) and pop (7; 21.9%) were the most frequent genres; 20 songs (62.5%) were international and 12 (37.5%) Brazilian. All songs had lyrics, and most were perceived as moderate in tempo (24; 75.0%) and energy (24; 75.0%) (Table 1). The full catalogue of songs is available in Online Resource 1.

Table 1 Characteristics of the personalized musical intervention (preferred songs chosen by the 32 participants)

Characteristic	n (%)
Different songs	31
Different artists	25
Genre	
Rock/alternative	9 (28.1)
Sertanejo	8 (25.0)
Pop	7 (21.9)
MPB	3 (9.4)
Rap/hip-hop	2 (6.3)
Indie/folk	2 (6.3)
Brazilian rock	1 (3.1)
Origin	
International	20 (62.5)
Brazilian	12 (37.5)
Lyrics	
Yes	32 (100.0)
Perceived tempo	
Moderate	24 (75.0)
Fast	6 (18.8)
Slow	2 (6.3)
Perceived energy	
Moderate	24 (75.0)
High	8 (25.0)

MPB, *música popular brasileira* (Brazilian popular music). Percentages refer to the 32 participants; one song was chosen by two participants. The full catalogue is provided in Online Resource 1

Performance by condition

Mean execution time and economy of motion were highest under the no-music condition in all three exercises, and lowest under random music in Pick and Place and Energy Dissection 1 and under

preferred music in Ring and Rail 1 (Table 2; Fig. 1). Because no music always corresponded to the first run and random music occurred only in the third run, these descriptive differences are confounded with run order and are consistent with the progressive reduction in execution time observed across consecutive runs. In the exploratory paired analysis, median execution time decreased from run 1 to run 2 in Pick and Place (52.78 vs 45.91 s; $p < 0.001$), Ring and Rail 1 (88.10 vs 73.11 s; $p = 0.003$) and Energy Dissection 1 (121.43 vs 91.74 s; $p < 0.001$), with 25, 24 and 28 of the 32 residents, respectively, performing faster in run 2.

Table 2 Simulator metrics by condition, pooling all runs performed under each condition

Metric	No music	Preferred music	Random music
Pick and Place			
Execution time (s)	58.45 (22.98)	48.96 (16.40)	44.33 (7.52)
	52.69 [30.84–138.41]	43.64 [25.99–92.47]	40.94 [36.07–58.49]
Economy of motion (cm)	165.63 (50.52)	157.31 (41.52)	152.43 (22.32)
	155.45 [103.32–354.22]	147.22 [104.12–302.24]	159.61 [123.36–185.94]
Workspace (cm)	9.08 (1.49)	9.13 (1.23)	8.94 (1.37)
	8.98 [6.85–14.17]	8.88 [7.32–11.91]	9.03 [7.26–11.69]
Instrument collisions (n)	0.69 (1.33)	0.43 (1.06)	0.00 (0.00)
	0.00 [0.00–6.00]	0.00 [0.00–4.00]	0.00 [0.00–0.00]
Instruments out of view (cm)	0.41 (2.68)	0.00 (0.00)	0.06 (0.22)
	0.00 [0.00–17.34]	0.00 [0.00–0.00]	0.00 [0.00–0.77]
Grasper contact warnings (s)	1.36 (3.60)	0.69 (1.69)	0.29 (0.54)
	0.00 [0.00–19.39]	0.00 [0.00–8.24]	0.00 [0.00–1.59]
Excessive instrument force (s)	0.11 (0.39)	0.07 (0.23)	0.04 (0.09)
	0.00 [0.00–2.07]	0.00 [0.00–1.35]	0.00 [0.00–0.27]
Drops (n)	0.43 (0.80)	0.38 (0.91)	0.25 (0.62)
	0.00 [0.00–3.00]	0.00 [0.00–5.00]	0.00 [0.00–2.00]
Ring and Rail 1			
Execution time (s)	101.34 (59.63)	74.21 (31.14)	79.51 (34.20)
	72.93 [29.74–264.25]	67.33 [32.92–172.96]	77.71 [41.51–153.78]
Economy of motion (cm)	152.28 (112.45)	119.22 (62.07)	133.26 (77.14)
	120.39 [33.11–616.22]	106.49 [46.46–330.11]	113.19 [61.73–314.97]
Workspace (cm)	9.59 (1.76)	8.96 (1.64)	8.52 (1.14)
	9.41 [7.03–15.51]	8.98 [5.56–14.01]	8.55 [6.91–10.70]
Instrument collisions (n)	1.41 (3.56)	0.73 (1.23)	1.00 (0.89)
	0.00 [0.00–18.00]	0.00 [0.00–5.00]	1.00 [0.00–2.00]

Metric	No music	Preferred music	Random music
Instruments out of view (cm)	7.29 (18.78) 1.35 [0.00–95.21]	3.00 (5.06) 0.34 [0.00–23.28]	2.80 (4.76) 0.06 [0.00–12.20]
Grasper contact warnings (s)	0.91 (2.67) 0.00 [0.00–15.57]	1.62 (3.28) 0.00 [0.00–15.79]	0.07 (0.20) 0.00 [0.00–0.67]
Excessive instrument force (s)	2.75 (9.00) 0.78 [0.00–55.46]	1.53 (4.62) 0.39 [0.00–30.07]	1.25 (1.82) 0.88 [0.00–6.00]
Drops (n)	0.44 (0.87) 0.00 [0.00–3.00]	0.18 (0.50) 0.00 [0.00–2.00]	0.00 (0.00) 0.00 [0.00–0.00]
Energy Dissection 1			
Execution time (s)	126.04 (45.90) 107.91 [62.58–249.60]	95.86 (48.99) 83.95 [55.83–295.96]	81.24 (32.93) 72.12 [45.96–148.88]
Economy of motion (cm)	179.49 (70.65) 176.07 [80.81–336.60]	151.56 (104.76) 120.97 [62.61–608.55]	124.90 (61.20) 96.55 [62.11–246.54]
Workspace (cm)	9.32 (2.16) 9.25 [4.72–17.45]	9.45 (2.48) 8.87 [5.90–16.92]	8.81 (2.51) 7.80 [6.16–14.63]
Instrument collisions (n)	1.38 (2.71) 0.00 [0.00–14.00]	0.79 (1.57) 0.00 [0.00–7.00]	0.85 (1.77) 0.00 [0.00–6.00]
Instruments out of view (cm)	3.96 (10.51) 0.00 [0.00–57.92]	9.65 (47.78) 0.00 [0.00–312.03]	2.54 (6.98) 0.00 [0.00–25.00]
Grasper contact warnings (s)	2.72 (6.90) 0.00 [0.00–35.25]	3.17 (6.63) 0.00 [0.00–27.24]	3.83 (9.78) 0.55 [0.00–35.88]
Excessive instrument force (s)	4.91 (11.63) 0.00 [0.00–45.10]	6.62 (22.91) 0.00 [0.00–141.01]	3.46 (9.29) 0.00 [0.00–34.00]
Inappropriately applied energy (s)	6.55 (5.29) 4.95 [0.00–18.87]	4.18 (4.34) 2.76 [0.00–21.92]	5.59 (6.07) 3.13 [0.00–21.44]
Blood loss (mL)	70.98 (91.68) 29.00 [0.00–333.90]	58.09 (139.67) 0.00 [0.00–762.90]	31.48 (44.36) 14.50 [0.00–141.30]

Metric	No music	Preferred music	Random music
Injured vessels (n)	0.33 (0.73)	0.35 (0.72)	0.46 (0.66)
	0.00 [0.00–3.00]	0.00 [0.00–3.00]	0.00 [0.00–2.00]

Values are mean (SD) and median [minimum–maximum]. No music always corresponded to run 1 and random music occurred only in run 3; differences between conditions are therefore confounded with run order. Each exercise pools run 1 (no music), run 2 (preferred music) and run 3 (condition drawn per exercise). Number of attempts (no music/preferred/random music): Pick and Place, 42/42/12; Ring and Rail 1, 41/44/11; Energy Dissection 1, 40/43/13. Units: s, seconds; cm, centimetres; mL, millilitres; n, count

Comparison between conditions in the randomized run

None of the 26 comparisons among the three conditions reached statistical significance after Bonferroni correction (Table 3). The lowest unadjusted p values were observed for instrument collisions in Pick and Place ($H = 6.19$, $p = 0.045$; corrected $p = 1.000$) and inappropriately applied energy in Energy Dissection 1 ($H = 7.08$, $p = 0.029$; corrected $p = 0.753$). Because no test was significant, no post hoc comparisons were performed. Effect sizes were small for most comparisons ($\epsilon^2 < 0.08$ in 20 of 26 tests) but were 0.08 or higher for six metrics, the largest being inappropriately applied energy in Energy Dissection 1 ($\epsilon^2 = 0.23$) and instrument collisions in Pick and Place ($\epsilon^2 = 0.20$). The distributions of the selected metrics overlapped considerably among the three conditions (Fig. 2).

Table 3 Kruskal–Wallis comparison of the three conditions in the randomly allocated run (run 3)

Exercise	Metric	H	df	p	p (Bonferroni)	ϵ^2
Pick and Place	Execution time (s)	0.73	2	0.693	1.000	0.024
	Economy of motion (cm)	0.26	2	0.880	1.000	0.008
	Workspace (cm)	0.04	2	0.982	1.000	0.001
	Grasper contact warnings (s)	1.32	2	0.516	1.000	0.043
	Instrument collisions (n)	6.19	2	0.045	1.000	0.200
	Excessive instrument force (s)	0.41	2	0.814	1.000	0.013
	Instruments out of view (cm)	1.67	2	0.435	1.000	0.054
	Drops (n)	0.59	2	0.743	1.000	0.019
Ring and Rail 1	Execution time (s)	2.35	2	0.309	1.000	0.076
	Economy of motion (cm)	1.86	2	0.395	1.000	0.060
	Workspace (cm)	1.87	2	0.394	1.000	0.060
	Grasper contact warnings (s)	2.30	2	0.317	1.000	0.074
	Instrument collisions (n)	4.76	2	0.092	1.000	0.154
	Excessive instrument force (s)	0.80	2	0.669	1.000	0.026
	Instruments out of view (cm)	0.29	2	0.867	1.000	0.009
	Drops (n)	2.56	2	0.279	1.000	0.082
Energy Dissection 1	Execution time (s)	1.69	2	0.430	1.000	0.054
	Economy of motion (cm)	0.44	2	0.802	1.000	0.014
	Workspace (cm)	2.01	2	0.366	1.000	0.065
	Grasper contact warnings (s)	1.80	2	0.407	1.000	0.058
	Instrument collisions (n)	1.99	2	0.369	1.000	0.064
	Excessive instrument force (s)	0.06	2	0.971	1.000	0.002
	Instruments out of view (cm)	0.10	2	0.953	1.000	0.003
	Inappropriately applied energy (s)	7.08	2	0.029	0.753	0.229
	Blood loss (mL)	4.06	2	0.131	1.000	0.131
	Injured vessels (n)	4.02	2	0.134	1.000	0.130

Attempts per condition (no music/preferred/random music): Pick and Place, 10/10/12; Ring and Rail 1, 9/12/11; Energy Dissection 1, 8/11/13. df, degrees of freedom; ϵ^2 , epsilon squared ($H/(n - 1)$), $n = 32$ per exercise. Corrected p values were obtained by multiplying the unadjusted p value by the number of tests (26) and truncating at 1.000

Discussion

In this study, we evaluated the influence of different musical conditions on the technical performance of general surgery residents during robotic simulator training. Unlike part of the literature, which investigated surgical performance using standardized music or genres defined by the researchers, such as classical, arousing or calming music [11, 19, 20], we included a condition in which each participant trained while listening to their own preferred music. Despite this approach, no statistically significant differences were observed among no music, preferred music and random music after correction for multiple comparisons.

These results are in line with part of the available literature. Systematic reviews show that, although some studies have identified favourable effects of music on certain components of surgical performance, the findings remain heterogeneous and do not establish a uniform benefit of musical intervention. Differences in experimental design, participant experience, task type and complexity, characteristics of the musical stimulus and exposure conditions are potential sources of this variability [7, 21] and limit direct comparison between studies.

Among studies with favourable results, preferred music has received particular attention. In laparoscopic simulation, music chosen by the participants was associated with significantly shorter execution time, shorter instrument path length and lower mental workload [8]. In robotic surgery, differences in performance according to the type of music were also observed, with shorter execution time and instrument path length under certain musical conditions [5]. However, these effects were not uniform across stimuli and tasks, suggesting that musical characteristics, individual preference and task complexity may modulate the influence of music on performance [5, 8].

Conversely, other experimental studies have not demonstrated a consistent technical benefit of music. In a virtual reality laparoscopic environment, no significant improvement attributable to music was observed, and changes in performance were related mainly to task repetition; arousing music was even associated with worse performance in the initial assessment, suggesting that musical stimulation may act as a distractor during the acquisition of new skills [11]. More recent results have also shown a lack of consistent benefit of music in certain simulation scenarios [13]. Notably, in a randomized crossover trial with 97 novices, preferred music reduced mental workload but did not improve laparoscopic task performance [14]. These findings agree with ours and reinforce that possible psychological or physiological effects of music do not necessarily translate into measurable improvement in objective performance metrics.

The detailed characterization of the musical intervention allows a broader interpretation of these results. The 32 participants chose 31 different songs by 25 artists, spanning genres such as rock, sertanejo, pop, *música popular brasileira* (MPB) and rap. This diversity shows that the intervention was based on individual preference rather than on the characteristics of a single genre, in contrast with most previous studies, which used standardized interventions, often classical music, while fewer evaluated songs chosen by the participants [7, 8]. Although preferred music has been associated with reduced mental workload and situational stress during simulated surgical tasks [8], in the present study this strategy did not translate into a significant improvement in technical performance. If music does influence performance in robotic simulation, its effect probably depends on the interaction of multiple factors rather than solely on exposure to a song the participant finds pleasant [5, 8].

The temporal pattern across runs is also relevant. Although the randomized comparison showed no differences among conditions, execution time decreased significantly from run 1 to run 2 in all three exercises, and most residents were faster in the second run. Because run 2 always followed run 1 and also introduced preferred music, this comparison cannot separate the two influences; however, the absence of a music effect in the randomized run suggests that the improvement is mainly explained by repetition. This pattern is consistent with a learning effect related to task repetition, previously described in simulated surgical training [11, 12], and indicates that part of the improvement across runs reflects progressive familiarization with the platform and the tasks, independently of the musical condition.

The composition of the sample should also be considered. All participants were residents without previous robotic experience, which provided homogeneity in baseline experience and reduced the influence of previously acquired robotic skills on the metrics. Previous investigations have similarly restricted their samples to novices or adopted criteria to minimize the influence of different skill levels [5, 7, 11]. In addition, the use of a commercial simulator that provides objective, standardized metrics reduced reliance on subjective assessment. The RobotiX Mentor has been validated for the assessment and training of robotic skills [1, 2], and quantitative parameters from simulators or motion analysis are widely used to assess simulated surgical performance [7].

Limitations

This study has limitations. First, only the third run was randomized; because no music always preceded preferred music, comparisons involving the first two runs are confounded by learning, and the descriptive differences between conditions should not be interpreted as effects of music. Second, the randomized comparison relied on small subgroups of 8 to 13 attempts per condition in each exercise, without an a priori sample size calculation, so the study was underpowered to detect small effects, and the moderate effect sizes observed for some metrics indicate that differences of practical relevance cannot be excluded; the Bonferroni correction across 26 tests, although protective against false-positive findings, further reduced power. Third, subjective outcomes such as anxiety, cognitive load, mental fatigue and perceived performance were not assessed, so psychological effects of music unaccompanied by changes in technical metrics cannot be excluded. Fourth, apart from the absence of previous robotic

experience, participant characteristics such as postgraduate year, video game experience or musical training were not recorded, so their potential influence on performance could not be assessed. In addition, music volume was not standardized, and familiarity with the random songs was not assessed, so some participants may have known the song they received. Finally, this was a single-centre study of novices performing basic tasks, which limits generalization to experienced surgeons and complex procedures. Nevertheless, the homogeneity of the sample, the use of objective metrics and the detailed characterization of the musical intervention lend consistency to the study and facilitate reproduction of the protocol.

Future studies should use fully randomized or counterbalanced sequences of conditions, a preparation phase to stabilize the initial learning curve before testing [14], larger samples, participants with different levels of experience, tasks of varying complexity and simultaneous assessment of technical, psychological and physiological parameters. This would make it possible to determine whether subjective effects of music, such as reduced stress or mental workload, occur in the absence of detectable changes in technical performance, and which individual, musical or task-related characteristics modulate this response.

Conclusion

Under the conditions evaluated, a personalized musical intervention was not associated with a statistically significant improvement in technical performance during robotic simulator training among general surgery residents. The reduction in execution time between consecutive runs is best explained by a learning effect. Personalized musical interventions proved feasible and can be characterized in detail, providing a methodological basis for larger, fully randomized studies integrating technical, physiological and cognitive outcomes.

Declarations

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Ethics approval

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Research Ethics Committee of the Faculty of Medicine of São José do Rio Preto (FAMERP) (CAAE 86407625.8.0000.5415; approval no. 7.531.150, 28 April 2025).

Consent to participate

Written informed consent was obtained from all individual participants included in the study.

Consent to publish

Not applicable. The manuscript contains no identifiable individual data or images.

Funding

No funding was received for conducting this study. Publication costs were supported by the Centro de Apoio à Pesquisa e Publicação (CENAP), Hospital de Base, FUNFARME/FAMERP.

Author Contribution

G.C.D. and G.G.S. conceived the study, contributed to its design, collected the data and wrote the original draft of the manuscript. C.D.S.C. contributed to the study design. R.J.G. and P.H.A.S. collected and organized the data. R.T.S. and M.A.R.F. performed the statistical analysis. H.Z.A. performed the scientific review and translation of the manuscript. G.G.S. supervised the study. All authors critically reviewed the manuscript and read and approved the final version.

Acknowledgement

The authors thank the Centro de Apoio à Pesquisa e Publicação (CENAP) of Hospital de Base, Fundação Faculdade Regional de Medicina de São José do Rio Preto (FUNFARME), and the Faculty of Medicine of São José do Rio Preto (FAMERP) for the final review of the manuscript and for supporting the publication costs, and the Hospital de Base for providing the robotic surgery simulator used in this study. The authors also thank the residents who participated in the study.

Data Availability

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

References

1. Whittaker G, Aydin A, Raison N, Kum F, Challacombe B, Khan MS, Dasgupta P, Ahmed K (2016) Validation of the RobotiX Mentor robotic surgery simulator. J Endourol 30:338–346. <https://doi.org/10.1089/end.2015.0620>

2. Leijte E, Claassen L, Arts E, de Blaauw I, Rosman C, Botden SMBI (2021) Training benchmarks based on validated composite scores for the RobotiX robot-assisted surgery simulator on basic tasks. *J Robot Surg* 15:69–79. <https://doi.org/10.1007/s11701-020-01080-9>
3. Allen K, Blascovich J (1994) Effects of music on cardiovascular reactivity among surgeons. *JAMA* 272:882–884
4. Narayanan A, Naidoo M, Kong V, Pearson L, Mani K, Fisher JP, Khashram M (2024) 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>
5. Siu KC et al (2010) The effect of music on robot-assisted laparoscopic surgical performance. *Surg Innov* 17:306–311
6. Lies SR, Zhang AY (2015) Prospective randomized study of the effect of music on the efficiency of surgical closures. *Aesthet Surg J* 35:858–863
7. Oomens P et al (2019) The effect of music on simulated surgical performance: a systematic review. *Surg Endosc* 33:2774–2784. <https://doi.org/10.1007/s00464-019-06868-x>
8. Oomens P, Fu VX, Kleinrensink VEE, Kleinrensink GJ, Jeekel J (2020) The effects of preferred music on laparoscopic surgical performance: a randomized crossover study. *World J Surg* 44:2614–2619. <https://doi.org/10.1007/s00268-020-05523-0>
9. Han Y, Zheng B, Zhao L, Hu J, Zhang C, Xiao R, Wang C, Pu D (2022) Impact of background music on the performance of laparoscopy teams. *BMC Med Educ* 22:439. <https://doi.org/10.1186/s12909-022-03503-7>
10. Moorthy K et al (2004) Objective evaluation of the effect of noise on the performance of a complex laparoscopic task. *Surgery* 136:25–30
11. Miskovic D et al (2008) Randomized controlled trial investigating the effect of music on the virtual reality laparoscopic learning performance of novice surgeons. *Surg Endosc* 22:2416–2420. <https://doi.org/10.1007/s00464-008-0040-8>
12. Shover A, Holloway J, Dauphine C, Benharash P, Xing H, Kansal N, Bowens N, Archie M, Kaji AH, de Virgilio C (2021) A randomized prospective blinded study evaluating the effect of music on novice surgical trainees' ability to perform a simulated surgical task. *J Surg Educ* 78:638–648. <https://doi.org/10.1016/j.jsurg.2020.08.020>
13. Narayanan A, Khashram M, Fisher JP (2025) Randomized cross-over trial comparing stress responses amongst undergraduates and surgeons with and without background music during simulated surgery. *Sci Rep* 15:19461. <https://doi.org/10.1038/s41598-025-02202-9>
14. Fu VX, Oomens P, Kleinrensink VEE, Sleurink KJ, Borst WM, Wessels PE, Lange JF, Kleinrensink GJ, Jeekel J (2021) 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. <https://doi.org/10.1007/s00464-020-07987-6>
15. Surgical Science (2026) RobotiX Mentor. <https://surgicalscience.com>. Accessed 30 June 2026

16. R Core Team (2024) R: a language and environment for statistical computing, version 4.4.1. R Foundation for Statistical Computing, Vienna. <https://www.R-project.org>
17. Nakazawa M (2024) fmsb: functions for medical statistics book with some demographic data. R package version 0.7.6. <https://CRAN.R-project.org/package=fmsb>
18. Wickham H (2016) ggplot2: elegant graphics for data analysis. Springer, New York. <https://doi.org/10.1007/978-3-319-24277-4>
19. Conrad C et al (2010) The effect of defined auditory conditions versus mental loading on the laparoscopic motor skill performance of experts. *Surg Endosc* 24:1347–1352
20. Wiseman MC (2013) The Mozart effect on task performance in a laparoscopic surgical simulator. *Surg Innov* 20:444–453. <https://doi.org/10.1177/1553350612462482>
21. El Boghdady M, Ewalds-Kvist BM (2020) 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>

Figures

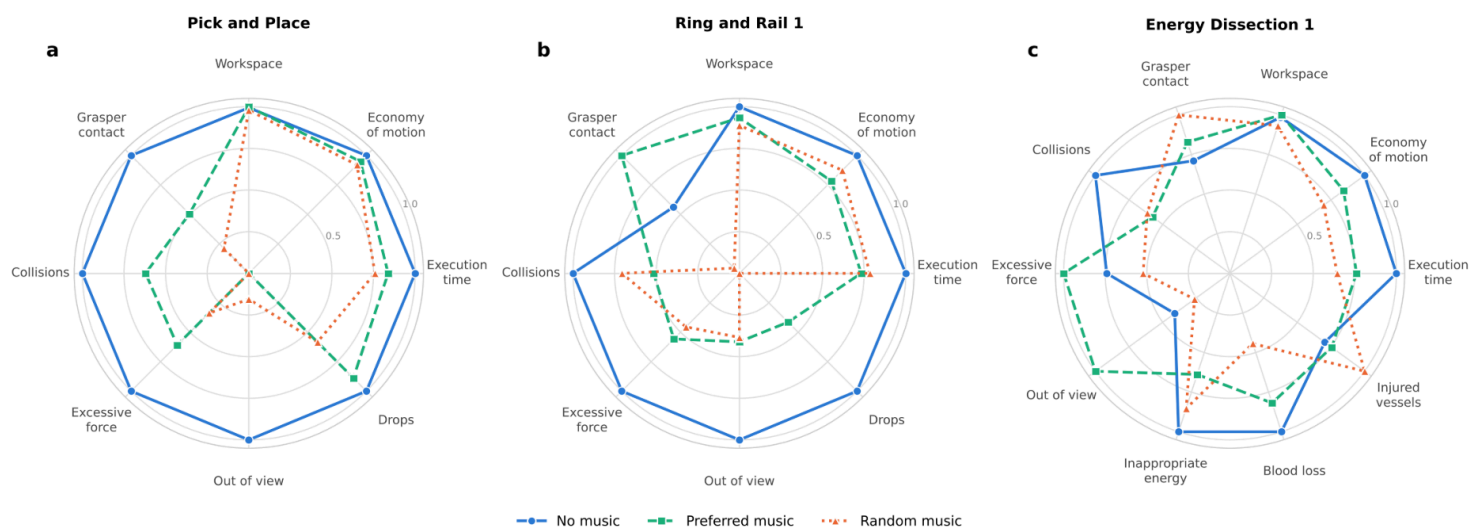


Figure 1

Radar charts of mean simulator metrics by condition, each metric scaled to the highest condition mean (= 1.0), for **a** Pick and Place, **b** Ring and Rail 1 and **c** Energy Dissection 1. Means pool all runs performed under each condition; because no music always corresponded to run 1 and random music only to run 3, differences are confounded with run order

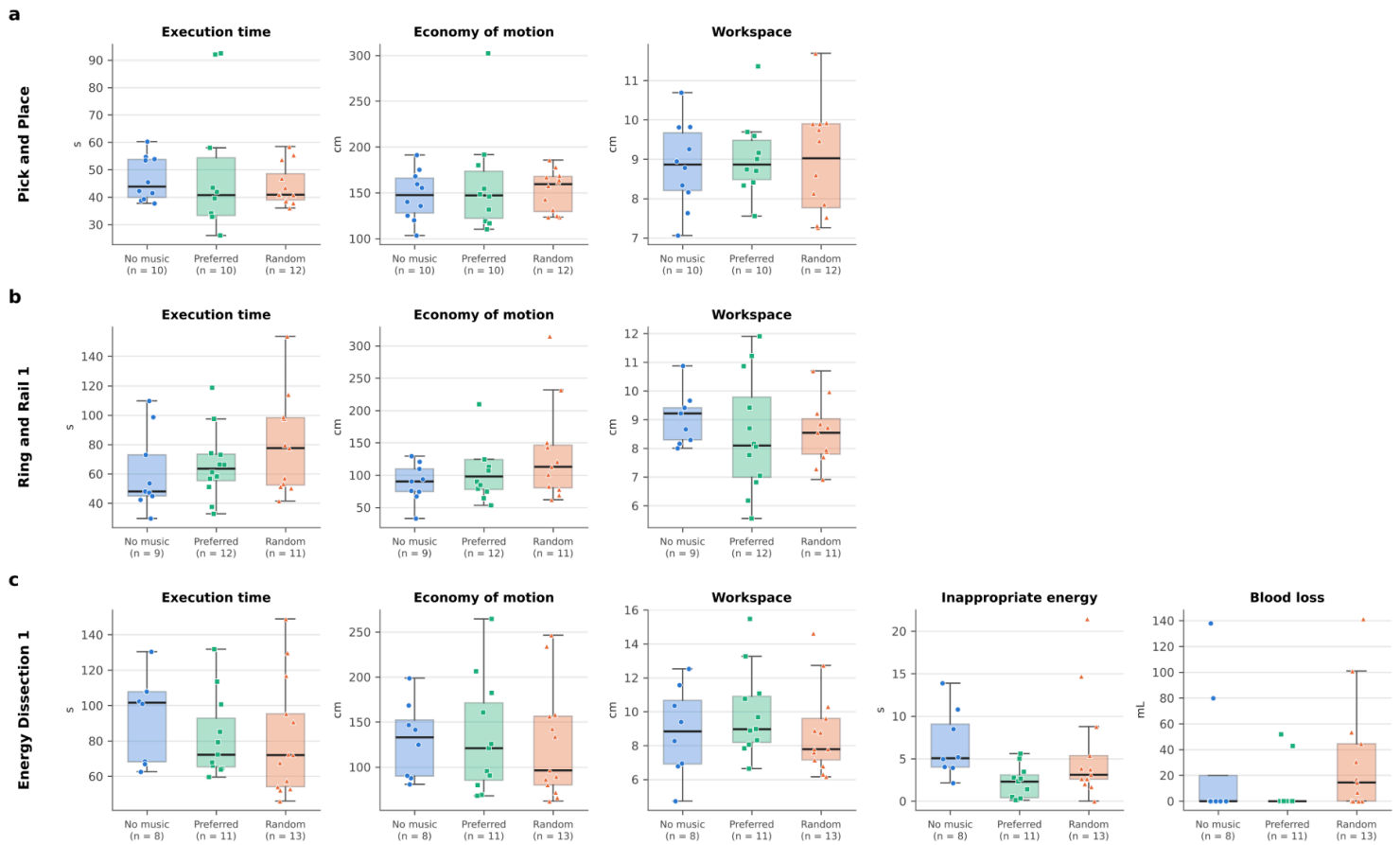


Figure 2

Distribution of selected simulator metrics by condition in the randomly allocated run (run 3). Points represent individual attempts; boxes show the median and interquartile range; whiskers extend to the most extreme values within 1.5 times the interquartile range