Mental Effort and Expressive Interaction in Expert and Student String Quartet Performance

Laura Bishop[1,2] (iD), **Simon Høffding**[1,3] (iD), **Olivier Lartillot**[1,2] (iD)
and Bruno Laeng[1,4]

Abstract

Resilience to changes in performance environments is a hallmark of expertise in music performance. Research has shown that skilled non-expert ensemble musicians maintain synchronization when visual interaction between them is disrupted, but that the quality of their expressive body motion changes. Our study extended these findings by testing how an expert string quartet responds to playing conditions that disrupt visual contact. We tested for potential effects on the musicians' expressive head motion, sound quality, and mental effort. The Danish String Quartet (DSQ), a world-class classical ensemble, performed an excerpt from a Haydn piece five times without an audience, then once for an audience of about 20 people. During the performances without audience, their seating configuration was manipulated to disrupt their audiovisual interaction. Audio, head motion, eye-tracking, and pupillometry data were collected. The DSQ's data were compared to data from a student quartet who completed the same experiment. Our results show that the DSQ maintained the quality of their musical sound and interactive body motion across disruptive and non-disruptive conditions, but mental effort (indexed by pupil size) was greater in non-disruptive conditions. In contrast, the student quartet moved less overall than the DSQ, moved less when they could not see each other, and did not show differences in pupil size across conditions. The quartets spent a similar percentage of performance time watching their co-performers. These findings suggest that the quality of audio and visual components of the DSQ's performance do not require visual interaction to maintain; however, these musicians do interact visually when given the opportunity. This visual interaction stimulates greater mental effort, perhaps reflecting increased social engagement.

Introduction

Expertise is associated with exceptional skill in a domain. Experts are fluent and flexible in their performance, carrying out skills consistently, accurately, and rapidly, but simultaneously adapting to variability in local conditions and novel task constraints (e.g., Lehmann et al., 2018). Experts in music performance, for example, can readily adjust their playing to accommodate different acoustic environments, spontaneously introduce new musical ideas, and recover from disruptions (Glowinski et al., 2016).

[1] RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion, University of Oslo, Oslo, Norway
[2] Department of Musicology, University of Oslo, Oslo, Norway
[3] Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark
[4] Department of Psychology, University of Oslo, Oslo, Norway

Corresponding author:
Laura Bishop, RITMO Centre for Interdisciplinary Studies in Rhythm, Time, and Motion, University of Oslo, Oslo, Norway.
Email: laura.bishop@imv.uio.no

Ensemble musicians must coordinate their sound across a range of expressive features, and with high precision, in addition to managing the technical and expressive aspects of playing their instrument in different environments (Bishop & Keller, 2022). To coordinate successfully, ensembles make use of real-time attention, anticipation, and adaptation processes (Keller et al., 2014). These processes require high-quality auditory (and to some extent visual) information flow between musicians (Bartlette et al., 2006; Bishop & Goebl, 2015).

How do ensemble musicians respond to playing conditions that stress or disrupt their established patterns of audiovisual interaction? In a recent study, a string quartet comprising advanced students became less communicative in their body motion when playing in conditions that prevented them from seeing one another directly (Bishop, González Sánchez, et al., 2021). The current study builds on this finding by replicating the same experiment with an expert string quartet. In addition, we introduce new evaluations of the effects of disrupted visual conditions on mental effort and sound quality.

Expressive Interaction in Diverse Musical Playing Conditions

In Western classical ensemble playing, a shared tempo and tight synchronization between players provide a foundation for expressive communication. Temporal synchronization is a multi-directional process in which all ensemble members continually adjust the timing of their actions (Konvalinka et al., 2010; Timmers et al., 2014). This is the case even if the ensemble has a clearly-defined leader: The followers adapt to the leader and the leader adapts to the followers, though leading/following relationships may emerge in note timing (Goebl & Palmer, 2009) and periodic body sway (Chang et al., 2017).

Although ensemble musicians mainly rely on audio signals to synchronize their playing, visual signals can also help. Musicians usually spend more time looking at the score than at their co-performers, but they do glance at their co-performers from time to time—especially at moments of uncertainty (e.g., when synchronizing after a long pause or during a tempo change; Kawase, 2014; Bishop et al., 2019a; Vandemoortele et al., 2018).

Patterns of visual attention between ensemble members relate to performer roles. In a case study of a student string quartet, the first violinist was watched more than any other musician but almost never looked at anyone else, perhaps reflecting his role as musical leader of the group (Bishop, González Sánchez, et al., 2021). In some cases, ensemble musicians may actively avoid watching their co-performers. This was observed in a study of singing duets, when the singers sang in canon rather than in unison (Palmer et al., 2019). Singers who were assigned a follower role turned away from their partner, presumably to avoid visual interference.

One of the emergent effects of visual information flow between co-performers is a strengthening of coordination in their ancillary body motion (Bishop et al., 2019b). For classical ensembles, this coordination tends to be stronger after rehearsing than before (Wood et al., 2022). Information flow in body sway, measured between musicians in a small ensemble using Granger causality, increases as the ensemble plays with more emotional intensity (Chang et al., 2019).

Coordination dynamics change when one ensemble member is given private instructions for how to play that are not shared with the rest of the group—simulating a situation where a performer veers from the practiced interpretation (Badino et al., 2014). The divergent player has less influence on their co-performers' head motion than when they play the music as practiced. In a recent study, a first violin section was repositioned so that they faced the second violin section and had their backs to the conductor. Intraperformer coordination between the head and bow increased for the repositioned violinists, and their head motion changed to more strongly demarcate the meter, suggesting that the violinists used their bodies to facilitate time-keeping (Laroche et al., 2022). Interpersonal coordination of head motion also increased for the repositioned violinists (Hilt et al., 2019).

In sum, despite the seeming secondary nature of the visual modality in ensemble performance, musicians communicate through body gestures when it might benefit the performance to do so. Coordination between musicians emerges in expressive body motion and changes according to the conditions surrounding the performance.

Mental Effort in Diverse Musical Playing Conditions

A recent study by Høffding et al. (2023) examined the effects of visual contact on shared absorption during string quartet playing, using cardiac synchrony as an index of shared absorption.[1] They observed negative effects of disruptive visual conditions for a student quartet, but not for an expert quartet. These negative effects suggest that, while students are less engaged during disruptive conditions, experts may be unaffected. The current study extends these findings by testing how disruptive visual conditions affect mental effort.

Mental effort refers to the amount of mental processing or arousal that is evoked in response to an event or task (Bruya & Tang, 2018; Kahneman, 1973). It is not equivalent to subjective effort and does not increase linearly in relation to subjective ratings of difficulty. Rather, it is an objective process that reflects a person's engagement with a task. In some contexts, mental effort seems to peak on tasks of moderate difficulty and decline for tasks that are easy or hard (Zekveld & Kramer, 2014).

Mental effort is commonly indexed through pupil size (Laeng & Alnaes, 2019). Increases in mental effort prompt pupil dilation (Kahneman, 1973). A distinction can be made between event-based (orienting) pupil responses and longer-duration responses that indicate tonic changes in mental processing and arousal (Mathôt, 2018).

Pupillometry has been used to show that vocal versions of melodies are more salient to listeners than instrumental versions and that familiar melodies elicit more attention than unfamiliar melodies (Weiss et al., 2016). Furthermore, recent studies have shown that listening to sounded music and imagining the same music evoke similar patterns of pupil dilation and constriction (Endestad et al., 2020; Kang & Banaji, 2020).

Pupillometry is rarely used in music performance settings, given that eye and body movements complicate interpretation of pupil data by introducing both measurement artifacts and attention demands. Nevertheless, a few studies have tested the effects of different performance demands on pupil size. O'Shea and Moran (2018) and Endestad et al. (2020) compared pupil responses in imagined versus overt performances. Bishop, Jensenius, et al. (2021) used pupillometry to test how the mental effort of performers and listeners fluctuated in response to technical, harmonic, and expressive demands in string quartet music. Expressive demands elicited pupil dilations from both performers and listeners, while technical demands elicited pupil dilations only for performers.

Overall, the literature shows that mental effort fluctuates across multiple timescales during music processing. For ensemble musicians, fluctuations in mental effort likely reflect processing of social information as well as processing of music. However, how mental effort varies in relation to the challenges and rewards of ensemble playing has not yet been investigated.

Current Study

The current study investigated how ensemble musicians respond to playing conditions that stress or disrupt their established patterns of audiovisual interaction. We addressed this question with a case study of an expert classical string quartet and tested three hypotheses. First, expert ensembles should be more resilient to varying performance conditions than student ensembles are. Therefore, we hypothesized that while students' body motion is less communicative when visual contact is disrupted (Bishop, González Sánchez, et al., 2021), experts' body motion is affected only in the most extreme conditions where no visual contact is possible at all. Second, we hypothesized that conditions that permit visual interaction between musicians are more engaging and evoke more mental effort, resulting in greater pupil dilation compared to conditions where musicians cannot interact visually. Third, we hypothesized that expert ensembles are generally more communicative in their body motion than students are.

An expert classical string quartet, the Danish String Quartet (DSQ), performed the same music five times without an audience, then once with an audience of about 20 people. During the first five performances, different seating configurations were used to manipulate whether the musicians could see and hear their co-performers normally. The quartet also sight-read a new piece that none of the musicians had heard or played before. We collected audio recordings, body motion data, and eye-tracking and pupillometry data as the quartet performed.

This study builds on a previous study that we carried out with a student group, the Borealis String Quartet ("BSQ"; Bishop, González Sánchez et al., 2021). We investigate how the DSQ's response to audiovisual manipulations differs from the response of the BSQ. We use some of the same analyses that we used for the student dataset, which we now apply to the experts' data, and introduce some new analyses, which we apply to both the experts' and students' data.

Methods

Participants

Participating in this study was the DSQ, a world-class professional ensemble that tours internationally and has recorded more than 10 albums together. The quartet in its current configuration was established in 2008, although three of the quartet members—the violinists and violist—have been playing chamber music together for more than 20 years. All members of the quartet provided written informed consent.

Design

Since no randomization was possible in the confines of this case study, performances were instead carried out in a fixed order, as described in Bishop, González Sánchez, et al. (2021).

Equipment and Materials

The quartet played bars 1–68 of Haydn's String Quartet in B-flat major, Op. 76, No. 4, *Allegro con spirito* for the first five performances, and for the final (concert) performance, they played the full movement. For the sight-read performance, they played Rued Langgaard's String Quartet No. 5, 2nd movement.

Head motion data were recorded using an OptiTrack motion capture system with 12 Flex 13 cameras, recording at 120 Hz. Body motion (shoulders, backs, arms, hands) was also captured but not analyzed here. Eye-tracking and pupillometry data were recorded using Pupil Core headsets from Pupil Labs, which were connected by cable to separate Hewlett Packard laptops running Pupil Capture software and recording at 200 Hz. DPA 4060 microphones were placed under the bridge of each instrument for audio recording. These were connected to a Behringer UMC404HD multichannel soundcard, which sent four-channel data to a MacBook Pro running Logic Pro. The musicians also wore wireless Delsys ECG biofeedback sensors for recording of heart activity. Heart activity data are reported on by Høffding et al. (2023).

A reflective marker was placed on a clapboard, which was struck once at the start of each recording to provide an audiovisual signal that was captured by the OptiTrack system and the audio recordings. Pupil Core recordings

were synchronized using the Pupil Groups and Time Sync plugins in Pupil Capture, which allow for time-synching of recording start and stop times for computers that are on a shared network. A Python script was run on one of the computers running Pupil Capture that documented the timestamps of incoming TTL triggers, which were sent by the OptiTrack system and received via Arduino. These triggers allowed retrospective synchronization between motion capture and eye-tracking/pupillometry data.

Procedure

The quartet spent a few minutes playing through the Haydn piece, which they had previously performed but had not played recently, and some other repertoire. They were then fitted with ECG sensors, markers for motion capture, and Pupil Core headsets. The calibration display from Pupil Capture was projected onto a large screen, and the musicians stood about 1.5 meters away as their Pupil Core devices were calibrated.

The quartet performed the 68-bar Haydn excerpt once in each of the following audiovisual (AV) conditions: Blind, Score-directed, Normal-rehearsal, Violin-isolated, and Replication-rehearsal (Figure 1; see descriptions below). They then recorded a single sight-read performance of the piece by Langgaard. Finally, they played the full movement of the Haydn for the Concert condition, for an audience of about 20 people.

- **Blind:** The musicians sat back-to-back so that they could not see each other at all and heard their co-performers from a different direction than they were used to.
- **Score-directed:** The musicians faced each other as normal but were instructed to look only at their scores, so could see their co-performers only peripherally.
- **Normal-rehearsal:** The musicians faced each other and had no restrictions on where they could look.
- **Violin-isolated:** The first violinist sat behind a curtain that visually occluded him from his co-performers and slightly changed his acoustic relationship with the others.
- **Replication-rehearsal:** Replication of Normal-rehearsal.
- **Sight-reading:** The musicians faced each other as normal and had no restrictions on where they could look as they sight-read an unfamiliar piece together.
- **Concert:** The quartet performed for a small audience while facing each other normally with no visual restrictions.

Analysis

The analyses described below were carried out on data from the DSQ and on data from the BSQ, collected in 2019 and reported on by Bishop, González Sánchez, et al. (2021) and Bishop, Jensenius, et al. (2021). Analyses of motion similarity, motion coupling, and pupil size per AV condition are new here for the BSQ dataset.
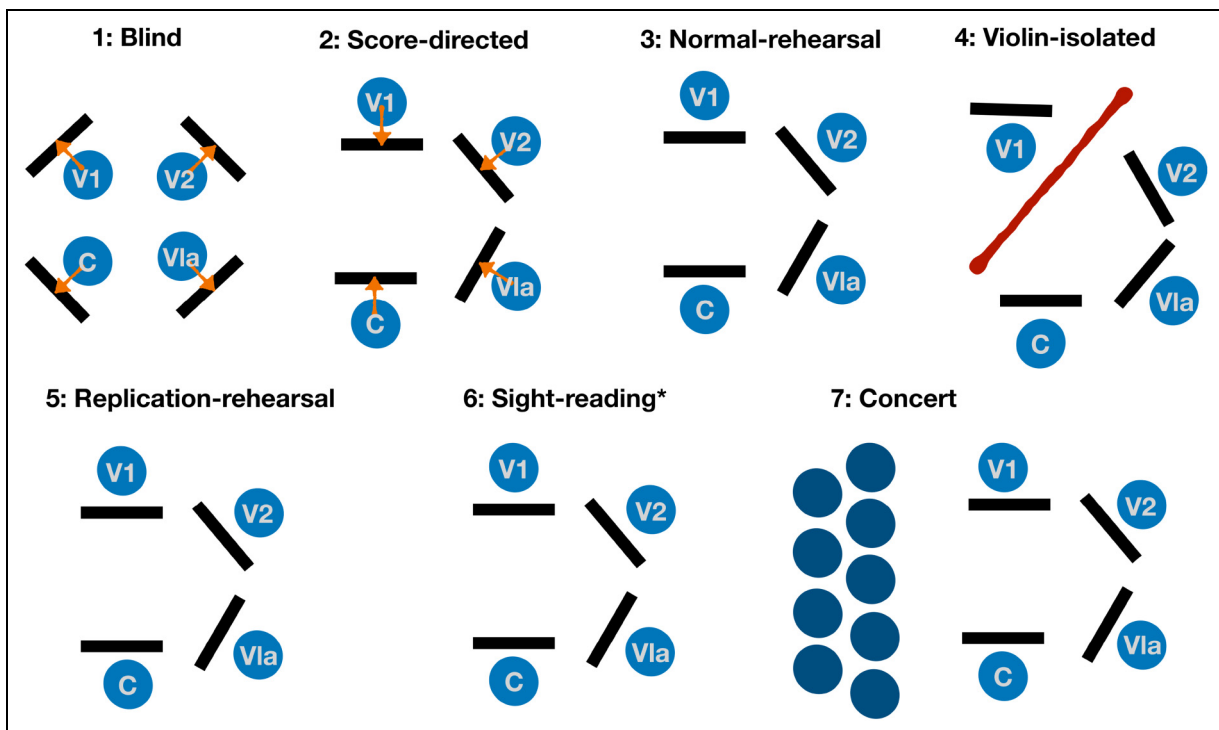

Figure 1. Audiovisual conditions that the DSQ performed, in order of completion. The Sight-reading condition (*) was new for the DSQ. All other conditions were previously tested in an experiment with the BSQ, a student quartet.

Head Motion Data. A spline interpolation was used to fill gaps in position data for the musicians' head markers. Position data were smoothed, and velocities were derived using a Savitzky–Golay filter ("prospectr" package in R; Stevens & Ramirez-Lopez, 2022). The Euclidian norms of three-dimensional values were then calculated. Three motion features were computed from these velocity norms: quantity of motion, between-performer motion similarity, and between-performer strength of coupling in motion periodicities. For a visual example, Figure 2
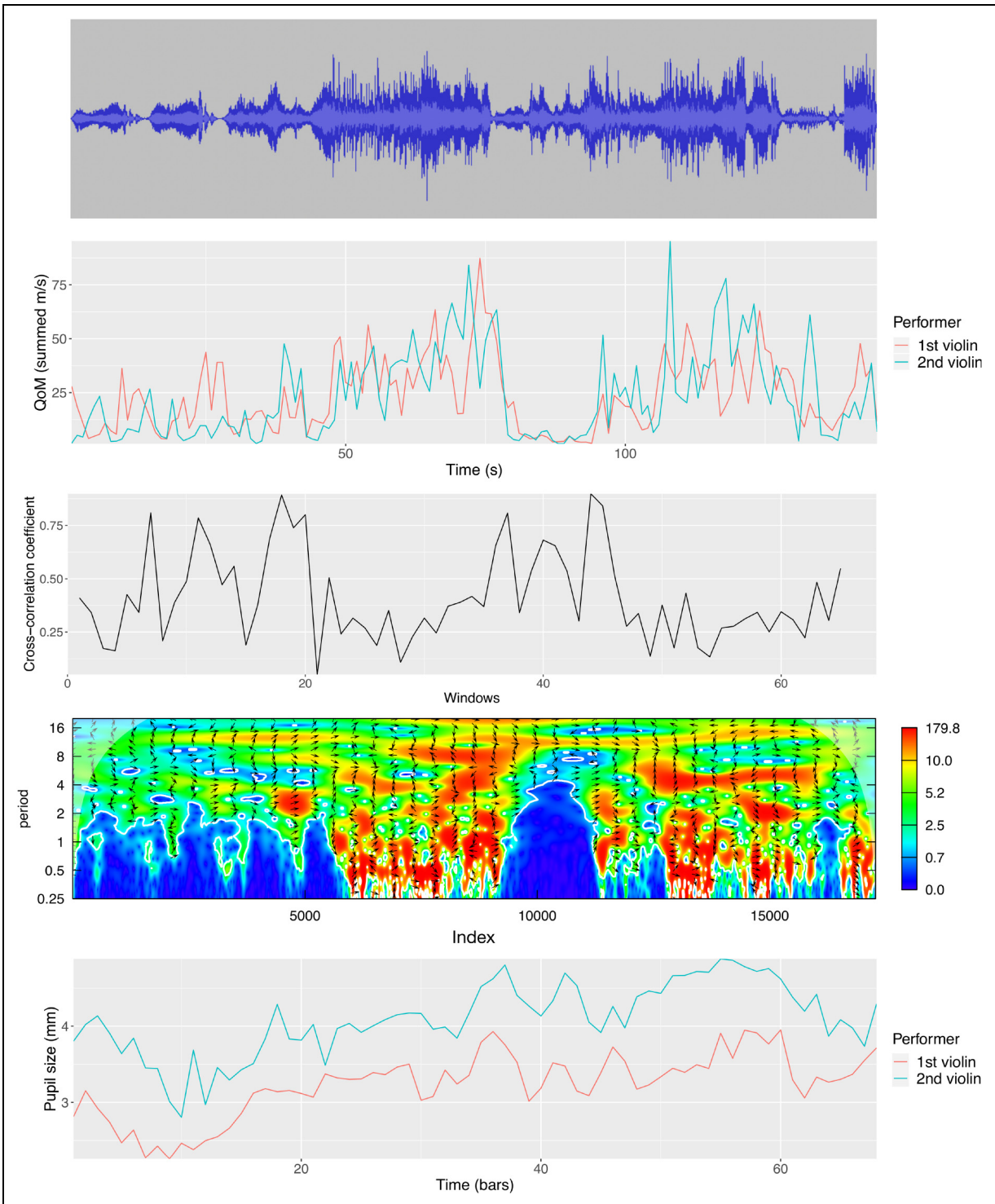

Figure 2. DSQ's performance in the Normal-rehearsal condition. From top: the audio waveform for the full quartet; QoM values per second for the violinists; cross-correlation coefficients per window for the comparison between violinists; the CWT power spectrum for the comparison between violinists; the pupil size curves for the violinists.

illustrates the time course of these features for the DSQ's 1st and 2nd violinists, for the Normal-rehearsal performance.

- **Quantity of motion (QoM)** was calculated as the sum of velocities per second. QoM values were highly skewed, so a log transformation was applied.
- **Motion Similarity** was calculated using cross-correlations, which were computed between each pair of musicians' head velocities using a rolling window, with a window size of two bars (i.e., 8 beats) and 50% overlap. A rolling window method was chosen to account for the possibility that the musicians might move more similarly in some parts of a performance than in others. The largest absolute correlation coefficient within each window at a lag of up to $+/-2$ beats was extracted, and these were averaged per pair and performance.
- **Motion Coupling** between performers in head motion frequencies was calculated using cross-wavelet transform analysis ("WaveletComp" package in R; Roesch & Schmidbauer, 2018). The analysis was run on a broad range of frequencies ranging from 0.05 to 4 Hz (i.e., periods between half a beat and 10 bars in duration, at the DSQ's average tempo). Power spectrums were then extracted for periods approximately equivalent to 1 bar ($+/-$ 1 beat), 2 bars ($+/-$ 2 beats), 4 bars ($+/-$ 3 beats), and 6 bars ($+/-$ 3 beats). The length of a phrase in the Haydn excerpt is mostly 6 bars (though this varies), and the length of a phrase in the Langgaard piece is 4 bars. We tested for differences in power between bar levels using linear mixed effects models (LMMs; "nlme" package in R; Pinheiro, et al., Pinheiro et al., 2021). The models included condition, musician pairing, and bar level as crossed random effects. Separate models were run for the BSQ's and the DSQ's performances of the Haydn piece and the DSQ's performance of the Langgaard piece. For the BSQ's Haydn performances, power was higher at 4 and 6 bars than at 1 and 2 bars ($p < .008$). For the DSQ's Haydn performances, power was higher at 6 bars than at all other bar levels ($p < .008$). For the DSQ's Langgaard performance, power was higher at 4 bars than at 1 and 2 bars. These results showed that power was strongest in the 4- and 6-bar levels across ensembles and pieces. Therefore, for the next stage of analysis, power values at the 4- and 6-bar levels were combined.

We compared the two quartets on each motion feature using LMMs. A separate model was run for each feature. These models included performer (for QoM) or performer pairing per quartet (for Motion Similarity and Motion Coupling) crossed with condition as random effects.

The effects of condition on motion features were also assessed using LMM, with a separate model run for each feature. These models included performer or performer pairing and condition as crossed random effects. For all models, the Normal-rehearsal condition was used as the baseline against which all other conditions were compared.

Pupillometry Data. Pupillometry data were filtered using a custom function developed in R. This function eliminated full and partial blinks based on: (1) absolute pupil values (i.e., pupil sizes of 0 were removed), (2) velocities in the pupil trajectory that exceeded a relative threshold (more than two standard deviations from the mean velocity for the trial), and (3) pupil values that exceeded a relative threshold (more than 2.5 standard deviations from the mean pupil diameter for the trial). The filtering function created some gaps in pupil curves that were filled using linear interpolation. Following interpolation, a Savitzky–Golay filter ("prospectr" package in R) was applied to smooth the curves. The quality of pupil tracking was estimated by calculating the proportion of samples that were excluded by the filtering function (prior to interpolation). Data from the best-tracked eye were retained for use in the analyses. Figure 2 shows pupil sizes, computed per bar, for the DSQ's 1st and 2nd violinist during the Normal-rehearsal. The cellist was dropped from all pupillometry analyses due to the poor quality of his data (he had worn glasses under the eye-tracking headset).

We tested the effects of AV condition on **Pupil Size,** averaged per performer and performance, using LMMs. The models included performer and condition as crossed random effects.

Eye-Tracking Data. Eye-tracking data were manually annotated using the video interface in Pupil Player (processing software from Pupil Labs), which shows a visual gaze position marker superimposed on footage from the eye-tracking headset's world-view camera, and allows the user to scroll forwards and backwards in time in one-frame increments. The temporal resolution of gaze annotations was the same as for the world-view camera (30 Hz). Gaze annotations were made for the DSQ in the three conditions where they played the Haydn excerpt and could watch each other freely (Normal-rehearsal, Replication-rehearsal, Concert). The cellist's gaze positions contained a lot of jitter and were not annotated.

Five areas of interest were defined for each performer: the score, their own bodies or instruments, and their (three) co-performers. We calculated the percentage of performance time that performers' gaze focused on each area of interest. *T*-tests were carried out to examine whether the DSQ watched their co-performers for different percentages of time across conditions. Percentages were subjected to an arcsine transformation before running the t-tests. Significance was evaluated at $\alpha = .02$ following Bonferroni correction.

An additional *t*-test was run to examine whether the BSQ spent a larger percentage of performance time watching their co-performers than did the DSQ. Since gaze was not

analyzed for the DSQ cellist, we also omitted gaze data for the BSQ cellist. Percentages were subjected to an arcsine transformation before running the *t*-test, and significance was evaluated at α = .05.

Audio. Audio recordings (captured per instrument) were processed using the version 1.9 of MIR toolbox (Lartillot et al., 2008) in MATLAB, and eight acoustic features were computed. These features included two measures of dynamics (short-term and longer-term), four measures of timbre (spectral centroid, spread, brightness, and roughness), and two measures of rhythm (pulse clarity, estimated using two methods based on maximum and on entropy; see Lartillot et al., 2008).

Audio features were extracted only for the Blind, Score-directed, Normal-rehearsal, Violin-isolated, and Concert performances—not the Replication-rehearsal, which we would not expect to differ much from the Normal-rehearsal. These are the same conditions that were subjected to listening evaluations (see Listening Evaluation experiment below).

We carried out two analyses on these feature data to determine whether any of the performance(s) showed a substantial acoustic difference from the Normal-rehearsal. The first analysis tested for differences at a global performance level. Feature data that failed a test of normality were log-transformed (brightness, spectral centroid, roughness, and spread required transformations). Data were then standardized per feature, and mean standardized values were calculated per feature, performer, and AV condition. LMM tested the effect of condition on feature values, with condition, performer, and feature included as crossed random effects to account for repeated measures. The Normal-rehearsal condition was set as the baseline level against which other conditions were compared. Significance was evaluated at α = .01 following Bonferroni correction.

The second analysis tested for differences between the Normal-rehearsal and other performances at a more fine-grained level. First, dynamic time warping was used to align log-filtered magnitude spectrograms for each recorded performance with the Normal-rehearsal (sampling rate 44.1 kHz, frame size 2,205 samples, hop size 441; Müller, 2021). This was done in Python. Audio feature data were resampled to match the timestamps of the aligned performances. The approximate timestamps of every second beat (i.e., 2 beats per bar) were then obtained manually in Sonic Visualizer, by tapping along with the Normal-rehearsal performance and inserting markers at taps. Audio feature data were averaged per feature, performer, performance, and 2-beat interval. Mean squared errors were then computed between the Normal-rehearsal performance and each of the Blind, Score-directed, Violin-isolated, and Concert performances. Finally, an LMM was run to test for an effect of condition on mean squared errors; condition, feature, and performer were included as crossed random effects.

As an indication of how timing varied within performances, aligned audio data were used to extract interbeat intervals (IBIs; between every second beat) in seconds for each performance. Since these timing data were obtained for the quartet rather than per performer, they could not be included in the analyses with the other acoustic features. Instead, we simply report the range and standard deviation of interbeat intervals per performance.

Results

Head Motion

We hypothesized that the DSQ's head motion would be less communicative in the most disruptive AV condition (i.e., Blind) compared to the Normal rehearsal. We also hypothesized that the DSQ would be generally more communicative in their head motion than the BSQ. Distributions of data for the motion features are shown in Figure 3, and the results of the models testing for effects of condition are listed in Table 1.

Overall, the DSQ moved more than the BSQ, β = 4.48, t(6) = 2.49, *p* < .05. AV condition had no significant effect on QoM for the DSQ. The BSQ moved less in the Blind and Score-directed conditions than in the Normal-rehearsal (as previously reported by Bishop, González Sánchez, et al., 2021).

Motion Similarity was higher for the DSQ than the BSQ, β = .02, t(10) = 2.23, *p* < .05. The DSQ showed lower Motion Similarity in the Replication-rehearsal and Sight-reading than in the Normal-rehearsal. For the BSQ, there were no effects of AV condition on Motion Similarity.

Motion Coupling did not differ in strength between the DSQ and BSQ, β = .03, t(10) = 1.03, *p* = .33. The DSQ showed higher Motion Coupling in the Blind and Sight-reading conditions than in the Normal-rehearsal, while the BSQ showed lower Motion Coupling in the Replication-rehearsal than in the Normal-rehearsal.

Pupillometry

We hypothesized that pupil size (indexing mental effort) would be larger in the conditions permitting visual interaction than in the conditions with visual disruption. Distributions of data are shown in Figure 3, and the results of the models testing for effects of condition are listed in Table 1. For the DSQ, Pupil Size was smaller in the Blind and Score-directed condition than in the Normal-rehearsal. No effects of AV condition on Pupil Size were observed for the BSQ.

Eye-Tracking

The percentages of performance time that the DSQ spent watching the score, their own bodies or instruments, and their co-performers in the Normal-rehearsal, Replication-
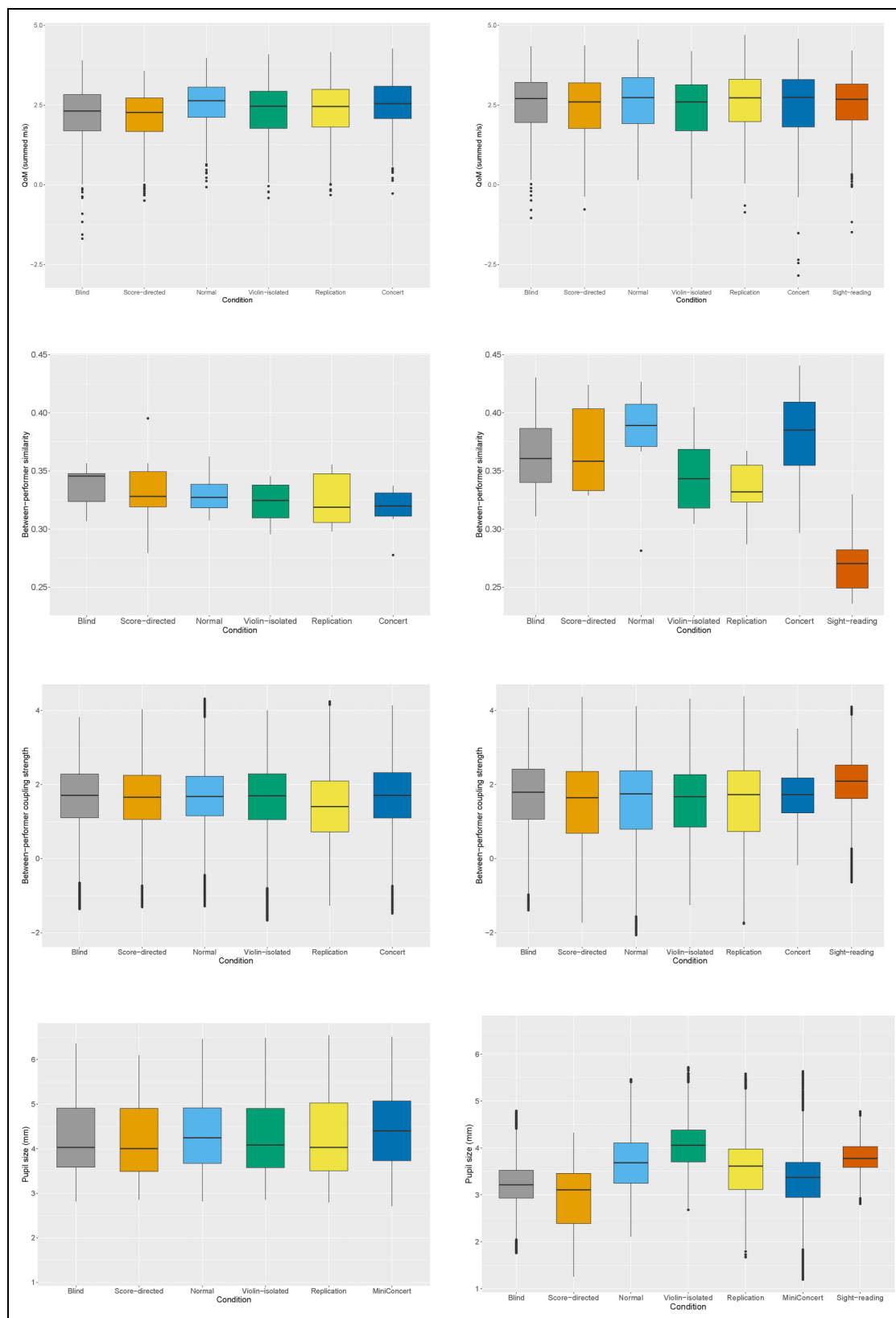
Figure 3. Violin plots showing, from top, QoM, Motion Similarity, Motion Coupling, and Pupil Size for the BSQ (left) and the DSQ (right).

Table 1. Results of LMMs testing the effects of audiovisual (AV) condition on head motion features and pupil size. Values show differences from Normal-rehearsal (β values from LMMs). Higher Motion Similarity and Motion Coupling values indicate stronger similarity/coupling. Note that the BSQ did not complete the sight-reading task.

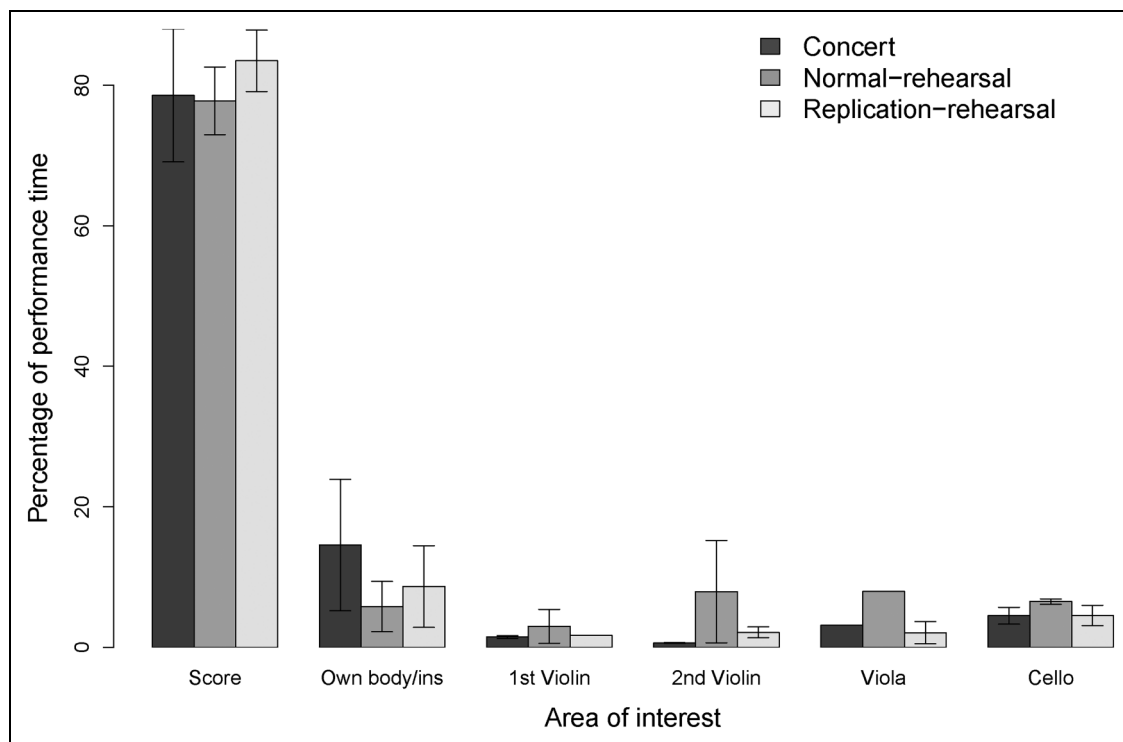|	QoM		Motion similarity		Motion coupling		Pupil size	
AV condition	DSQ	BSQ	DSQ	BSQ	DSQ	BSQ	DSQ	BSQ
Blind	−.08	−.36*	−.012	.006	.19**	−.02	−.49*	−.09
Score-directed	−.19	−.40*	−.009	.004	−.03	−.04	−.70**	−.17
Violin-isolated	−.25	−.22	−.030	−.007	.03	−.03	.33	−.05
Replication-rehearsal	−.04	−.23	−.044*	−.005	.02	−.27**	−.12	−.08
Concert	−.09	−.06	.001	−.014	.19	−.01	−.36	.10
Sight-reading	−.08	—	−.105**	—	.54**	—	.09	—

Note. *p < .008, **p < .001.

Figure 4. Percentages of performance time that the DSQ spent looking at the score, their own body or instrument, and their co-performers. Percentages have been averaged across performers. Error bars indicate standard error. Where error bars are missing, there were not enough data points to calculate standard error (i.e., only one performer looked at that area of interest during the performance).

rehearsal, and Concert performances are shown in Figure 4. *T*-tests showed no significant difference between these conditions in the percentages of time that the performers spent watching each other. A *t*-test comparing the percentages of performance time that the BSQ and DSQ musicians spent watching their co-performers showed no significant difference between ensembles (BSQ $M = 4.7\%$, $SD = 4.8\%$; DSQ $M = 3.9\%$, $SD = 3.4\%$). The percentages of performance time that each musician in the DSQ spent watching each other musician are given in Table 2.

Table 2. Percentages of performance time that the DSQ musicians spent watching each other. Note that gaze data were not analyzed for the cellist.

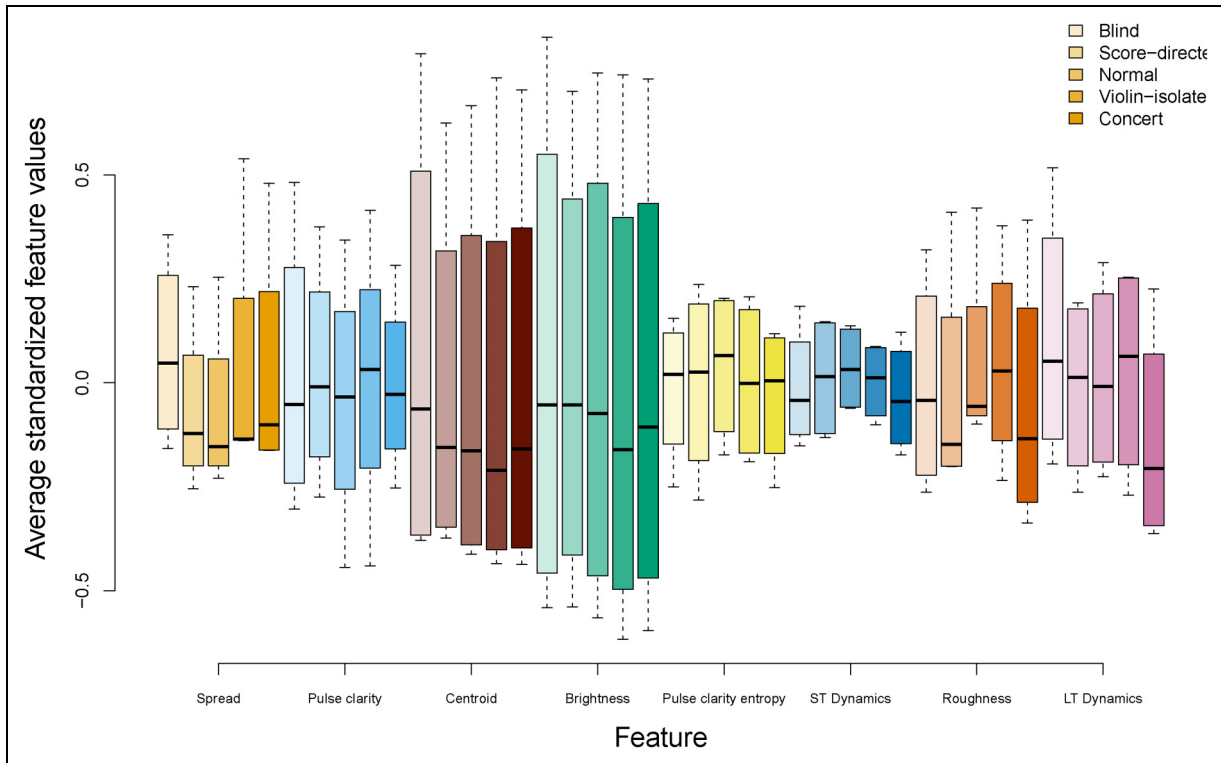|		Viewing musician		
		1st Violin	2nd Violin	Viola
Viewed musician	1st Violin	–	0.9%	2.9%
	2nd Violin	6.3%	–	0.8%
	Viola	4.9%	0.3%	–
	Cello	6.4%	5.8%	3.4%

Figure 5. Boxplot showing standardized acoustic feature values, averaged across DSQ musicians, for each condition. The conditions are listed in the same order from left (low saturation) to right (high saturation) for all features. ST and LT dynamics indicate short- and long-term dynamics.

Table 3. Ranges and standard deviations of interbeat intervals (computed every other beat; i.e., 2 IBIs per bar) across conditions.

Condition	Min	Max	Range (Max–Min)	Standard deviation
Blind	0.61	1.46	0.85	0.11
Score-directed	0.61	1.49	0.88	0.11
Normal-rehearsal	0.86	1.61	0.75	0.10
Violin-isolated	0.82	1.62	0.82	0.12
Concert	0.70	1.62	0.92	0.14

Audio

The LMM testing for an effect of AV condition on mean standardized feature values yielded no significant effects at $\alpha = .01$ (Figure 5). The LMM testing for an effect of condition on mean squared errors (indicating differences from the Normal-rehearsal performance) yielded a non-significant main effect of condition, $F(1,3) = 2.39$, $p = .07$. The range and standard deviation of interbeat intervals per performance are given in Table 3. Overall, the Concert performance tended to be more variable in timing, while the Normal-rehearsal performance tended to be less variable (though these differences could not be tested statistically given the small sample size).

Listening Evaluation Experiment

Our analysis of acoustic features in the DSQ's musical sound did not reveal any major differences between the performances given under disrupted AV conditions and the Normal-rehearsal performance. However, these acoustic analyses may not capture differences in musical quality that human listeners are sensitive to. Our listening evaluation experiment was intended to address this point.

Participants

A total of 189 listeners participated in the experiment (112 female, 69 male, 8 non-binary; age $M = 36.8$, $SD = 14.5$). Listeners were recruited from the DSQ FaceBook fan page ($n = 50$) and Prolific ($n = 139$). Participants from Prolific were required to have five or more years' experience playing the violin, viola, or cello. Participants self-identified as professional musicians (23), semi-professional musicians (27), serious amateur musicians (46), amateur musicians (72), music-loving nonmusicians (19), or nonmusicians (3). Asked to rate how often they listen to string quartet music on a five-point scale ranging from "never" to "very often," they gave an average rating of 3.0 ($SD = 1.0$).

Design

Participants were randomly assigned to one of eight conditions. In each condition, participants heard a recording of the DSQ's Normal performance and a recording of one of their experimental performances (Blind, Score-directed, Violin-isolated, or Concert). The order of normal and

Table 4. Counts of listeners' votes for their preferred performances. Each listener heard the Normal-rehearsal performance and one of the experimental performances. No significant differences arose between performances.

Performances compared	Number of votes
Normal-rehearsal	27
Blind	19
Normal-rehearsal	21
Score-directed	28
Normal-rehearsal	24
Violin-isolated	24
Normal-rehearsal	21
Concert	25

experimental performances was counterbalanced across conditions.

Procedure

The experiment was completed online. Listeners were presented with two audio recordings, one after the other, and indicated which recording they preferred. They were required to listen to both recordings in their entirety before stating their preference. No information about the context of the recordings or manipulation of AV conditions was provided. Listeners also answered a few demographics questions. The listening task was combined with another (audiovisual) performance evaluation task that involved other music performed by the DSQ, and which participants completed first.

Analysis and Results

Chi-squared tests were run to test for preferences between the Normal performance and each of the experimental performances (Table 4). No significant differences in preference counts were observed (all $p > .05$). As a follow-up, to see whether more experienced listeners had more consistent preferences for the Normal over experimental performances, we tested the correlation between participants' listening habit ratings (1–5; see "Participants") and a dummy-recoded preference value (1 if the Normal recording was chosen; 0 if the experimental recording was chosen). None of these correlations was significant (all $p > .05$; $-.11 < r < .19$). Thus, there was no reliable preference between Normal and experimental conditions, even when listening habits were accounted for.

A Chi-squared test also compared participants' preferences between the first and second recordings, independent of which performances they contained. This test revealed a significant preference for the second recording over the first, $X^2 = 5.76$, $p = .02$.

Discussion

This study investigated how expert string quartet musicians were affected by playing conditions that manipulated their ability to visually interact and the presence or absence of an audience. We examined the musicians' overt behavioral responses (in head motion and musical sound) and covert attentional responses (indexed by pupil size). An expert ensemble, the Danish String Quartet (DSQ), with more than a decade of experience playing together, was compared with a student ensemble, the Borealis String Quartet (BSQ), who had been playing together for less than a year.

The DSQ was largely resilient to changes in audiovisual conditions. Yet despite their lack of overt response, they showed smaller pupil sizes in the Blind and Score-directed conditions than in the Normal-rehearsal. This effect is in line with our hypothesis that visual interaction evokes increased mental processing. The DSQ moved more overall and were more coordinated than the BSQ, but DSQ and BSQ musicians spent similar amounts of time watching their co-performers. These results are discussed in more detail below.

Response to Disruptive Playing Conditions

We predicted that the DSQ would reduce their communicative body motion only in the Blind condition, which prevented visual interaction entirely. Instead, the DSQ proved more resilient than we expected, and moved a similar amount across all conditions. This lack of effect contrasts with that observed for the BSQ and student musicians in other studies (Bishop et al., 2019b), who move less when they cannot see their co-performer(s). The reduced QoM that we observe among students who play in visually isolating conditions might occur because these conditions encourage internally directed attention (to the player's own part). In contrast, conditions that allow for visual interaction might encourage externally directed attention (to the partner(s) and shared musical output). The DSQ, for whom QoM did not differ across conditions, might have stronger attention strategies that are less dependent on changes in playing conditions.

The negative effect of visual disruption on mental effort for the DSQ is a novel finding. This effect suggests that visual interaction influences players' experience of performing together even if it has little impact on the quality of their performance. One might question why ensemble musicians choose to watch each other while playing when processing that social-visual information seemingly requires more mental "work." Our interpretation is that this extra "work" is rewarding. Visual interaction may support the emotional connection between players and promote a sense of musical togetherness. The BSQ did not show the same differences in pupil size, perhaps because they were not as flexible in redistributing their attention across conditions as the DSQ were and/or were more affected by performance anxiety.

Contrary to our expectations, neither quartet showed a marked response to the Concert in terms of either body motion or pupil size. This lack of effect is less surprising for the DSQ, who are used to performing professionally

for audiences of hundreds of people, but more surprising for the BSQ, for whom the concert was a notable occasion (and also comprised their semester exam; see Bishop, Jensenius, et al., 2021). Prior research has shown that performers may strengthen the communicative features of their body motion in front of an audience, for example, moving with greater intensity (Moelants et al., 2012; Schaerlaeken et al., 2017). These effects may be tempered by factors such as performance anxiety coupled with an error-avoidant mindset, which might promote a more constrained style of expressive motion, or acclimatization to the presence of the audience.

Surprisingly, the DSQ showed reduced motion similarity and the BSQ showed reduced motion coupling in the Replication-rehearsal, compared to the Normal-rehearsal. We had not predicted that differences would arise between the Normal-rehearsal and Replication-rehearsal, given that the audiovisual conditions were the same. These effects show that even in conditions that allow for normal visual interaction between quartet co-performers, the strength of coordination in head motion can vary, possibly reflecting changes in the musicians' focus of attention, experimentation with different interpretations, or extramusical factors such as fatigue. Also surprising was the strengthened motion coupling that arose for the DSQ in the Blind condition. This strengthened coupling might have arisen as the quartet sought to maintain coordination under more challenging conditions. Strengthened coupling could reflect a more predictable style of expressive body motion that facilitates timekeeping and coordination.

The DSQ showed stronger motion coupling while sight-reading the Langgaard piece than during the Normal-rehearsal performance of the Haydn, perhaps reflecting a shared tendency to temporally align head motion with the music during the early stages of learning a new piece. Such a tendency might help with timekeeping. The strong coupling observed during the Sight-reading might also be attributable to structural and stylistic differences between the Langgaard and Haydn pieces. Despite the strengthened motion coupling, the DSQ showed reduced motion similarity while Sight-reading relative to the Normal-rehearsal. Motion similarity represents the degree of alignment between musicians' velocity curves within 2-bar windows, while motion coupling represents the strength of alignment in periodic motion elapsing over 4–6-bar intervals. Thus, these measures capture different aspects of motion quality. Differences between them show how important it is to use multiple measures of motion quality to get a comprehensive description of how ensemble musicians relate to each other through body motion.

Visual Interaction through Body Motion and Gaze in Expert and Student Quartets

We hypothesized that the DSQ would be more communicative in their expressive body motion than the BSQ. Our results partially support this hypothesis. The DSQ moved more overall and had stronger Motion Similarity than the BSQ, but the quartets did not differ in Motion Coupling. The DSQ's higher quantity of motion might be a marker of greater expressivity and higher-quality performance, or simply a stylistic choice of the ensemble. The literature on expressive music performance shows a relationship between quantity or intensity of body motion and the expressive intensity in musical sound. For example, musicians move substantially more when asked to play with exaggerated expressivity than when asked to play with minimal expressivity (Massie-Laberge et al., 2019; Thompson & Luck, 2012). However, expressive intensity does not necessarily share a positive linear relationship with performance quality, since high-quality performance usually involves stylistically appropriate expressive contrasts, or changes between expressively more intense (with larger-scale body motion) and less intense (with smaller-scale body motion) periods.

The DSQ's stronger Motion Similarity compared to the BSQ also raises questions about how coordination in expressive body motion between ensemble co-performers relates to performance quality. Information flow in body sway between ensemble members strengthens when the ensemble plays with more intense emotional expression (Chang et al., 2019). Audiences are sensitive to the strength of motion coordination between ensemble members (Jakubowski et al., 2020), and visually more-coordinated ensembles are judged as more "together" than visually less-coordinated ensembles (D'Amario et al., 2022). Thus, visual coordination of expressive motion may communicate some aspect of "togetherness." Further study is needed to determine whether coordination of expressive motion contributes in some way to coordination of musical sound.

The DSQ, like the BSQ (reported on in Bishop, González Sánchez, et al., 2021), spent most of their performance time watching the score. Both quartets spent approximately the same percentage of performance time watching their co-performers. Glances between co-performers reflect the social nature of ensemble performance and are evidence that ensemble interactions extend beyond the auditory domain. Further research is needed to clarify whether glances between co-performers enhance the social rewards of playing music with others or are simply a byproduct of musicians' changing focus of attention. To address this directly, the quality of ensemble musicians' experiences should be considered alongside measurement of their gaze patterns.

In our previous paper (Bishop, González Sánchez, et al., 2021), we reported on asymmetries that arose in the gaze patterns of the BSQ musicians: While the 1st violinist was watched more by his co-performers than any other ensemble member, he rarely looked at anyone else. We interpreted this as potentially reflecting the ensemble's view of the 1st violinist as the musical leader. The same asymmetries did not arise in the DSQ's gaze patterns. It is also interesting to note that instances of direct gaze seem to occur less frequently between musicians who are seated directly next to each other and more frequently

between musicians who are seated further apart. Direct gaze between musicians who are seated side-by-side might require rotation of the body and, therefore, more deliberate effort.

Limitations

Some limitations of this study should be noted. First, our study focused on two string quartets and their performance of a short excerpt from one piece of classical repertoire. Further research is needed to investigate whether the effects that we have observed generalize to other ensembles performing other repertoire.

Our analyses of head motion focused on low-level features and tested for linear effects of AV conditions. Future studies might consider higher-level motion features, such as fluidity, impulsivity, directness, and contraction (Camurri et al., 2004). Such measures might give a more nuanced understanding of how expressive motion is shaped in interactive settings. Future studies might also implement analyses that account for possible non-linear effects and group-level coordination patterns in addition to pairwise relationships (Demos & Palmer, 2023).

While the BSQ and DSQ performed the same music under the same manipulations, they did so in different contexts and with some differences in recording equipment. The BSQ completed the experiment in our laboratory, while the DSQ completed it in a concert hall. This difference was unavoidable, as due to scheduling challenges, it was not possible to bring the DSQ to our lab. This change in setting might have led to differences in atmosphere and experience for the musicians.

More issues with data quality arose for the DSQ, given the poorly controlled environment of the concert hall. In terms of recording equipment, a Qualisys motion capture system and eye-tracking glasses from SMI were used for data collection with the BSQ, while an OptiTrack motion capture system and eye-tracking headsets from Pupil Labs were used with the DSQ. Our prior research has shown that the Qualisys and OptiTrack systems are comparable in noise level (Bishop & Jensenius, 2020). We did note more noise in the eye-tracking/pupillometry data collected using Pupil Labs headsets, but this was likely due to the less controlled lighting conditions of the concert hall.

Conclusions

This study investigated the resilience of string quartet musicians to playing conditions that disrupt their practiced patterns of co-performer interaction. Our results show that expert musicians can maintain the quality of their musical sound and expressive body motion in disrupted conditions, but mental effort is greater in conditions that allow co-performers to interact visually. The student quartet, in contrast to the expert quartet, moved less in conditions that did not allow for visual interaction and showed no differences across conditions in mental effort. Musicians in both quartets spent a small percentage of the playing time watching their co-performers. Overall, our findings point to high flexibility in the processes involved in maintaining coordination and joint expressivity during skilled ensemble performance and suggest that for experts, visual interaction stimulates increased mental effort, perhaps associated with the processing of social information. Future research should focus on the potential effects that visual interaction has on the quality of social rewards that result from ensemble playing.

Action Editor

Alexander Refsum Jensenius, University of Oslo, RITMO Centre for Interdisciplinary Studies in Rhythm, Time and Motion.

Peer Review

Gualtiero Volpe, Universita degli Studi di Genova, Dipartimento di Informatica, Bioingegneria, Robotica, e Ingegneria dei Sistemi. Geoff Luck, University of Jyväskylä, Department of Music, Art, and Culture Studies.

Contributorship

LB, SH, and BL conceived and designed the study. LB and SH collected the data. LB and OL contributed to data analysis, and all authors contributed to interpreting the results. LB wrote the first draft of the manuscript, then all authors reviewed and edited the manuscript and approved the final version.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval

The Norwegian Centre for Research Data (NSD) has approved this study (reference no. 748915).

ORCID iDs

Laura Bishop https://orcid.org/0000-0002-0656-3969
Simon Høffding https://orcid.org/0000-0002-9739-9454

Olivier Lartillot 🆔 https://orcid.org/0000-0001-6129-686X

Note

1. These cardiac data were collected during the same recording sessions as we report on here.

References

Badino, L., D'Ausilio, A., Glowinski, D., Camurri, A., & Fadiga, L. (2014). Sensorimotor communication in professional quartets. *Neuropsychologia*, *55*, 98–104. https://doi.org/10.1016/j.neuropsychologia.2013.11.012

Bartlette, C., Headlam, D., Bocko, M., & Velikic, G. (2006). Effect of network latency on interactive musical performance. *Music Perception*, *24*(1), 49–62. https://doi.org/10.1525/mp.2006.24.1.49

Bishop, L., Cancino-Chacón, C., & Goebl, W. (2019a). Eye gaze as a means of giving and seeking information during musical interaction. *Consciousness and Cognition*, *68*, 73–96. https://doi.org/10.1016/j.concog.2019.01.002

Bishop, L., Cancino-Chacón, C., & Goebl, W. (2019b). Moving to communicate, moving to interact: Patterns of body motion in musical duo performance. *Music Perception*, *37*(1), 1–25. https://doi.org/10.1525/mp.2019.37.1.1

Bishop, L., & Goebl, W. (2015). When they listen and when they watch: Pianists' use of nonverbal audio and visual cues during duet performance. *Musicae Scientiae*, *19*(1), 84–110. https://doi.org/10.1177/1029864915570355

Bishop, L., González Sánchez, V., Laeng, B., Jensenius, A. R., & Høffding, S. (2021). Move like everyone is watching: Social context affects head motion and gaze in string quartet performance. *Journal of New Music Research*, *50*(4), 392–412. https://doi.org/10.1080/09298215.2021.1977338

Bishop, L., & Jensenius, A. R. (2020). Reliability of two IR motion capture systems in a music performance setting. In *Proceedings of the sound and music computing conference*, Torino, Italy.

Bishop, L., Jensenius, A. R., & Laeng, B. (2021). Musical and bodily predictors of mental effort in string quartet music: An ecological pupillometry study of performers and listeners. *Frontiers in Psychology*, *12*, 653021. https://doi.org/10.3389/fpsyg.2021.653021

Bishop, L., & Keller, P. E. (2022). Instrumental ensembles. In *The Oxford Handbook of Music Performance* (Vol. 2, pp. 417–441). Oxford University Press.

Bruya, B., & Tang, Y.-Y. (2018). Is attention really effort? Revisiting Daniel Kahneman's influential 1973 book attention and effort. *Frontiers in Psychology*, *9*, 1133. https://doi.org/10.3389/fpsyg.2018.01133

Camurri, A., Mazzarino, B., & Volpe, G. (2004). Analysis of expressive gesture: The eyesweb expressive gesture processing library. In A. Camurri & G. Volpe (Eds.), *Gesture-based communication in human-computer interaction* (pp. 460–467). Springer Berlin Heidelberg.

Chang, A., Kragness, H. E., Livingstone, S. R., Bosnyak, D. J., & Trainor, L. J. (2019). Body sway reflects joint emotional expression in music ensemble performance. *Scientific Reports*, *9*(1), E4134–E4141. https://doi.org/10.1038/s41598-018-36358-4

Chang, A., Livingstone, S. R., Bosnyak, D. J., & Trainor, L. J. (2017). Body sway reflects leadership in joint music performance. *Proceedings of the National Academy of Sciences*, *114*(21). https://doi.org/10.1073/pnas.1617657114

D'Amario, S., Goebl, W., & Bishop, L. (2022). Judgment of togetherness in performances by musical duos. *Frontiers in Psychology*, *13*, 997752. https://doi.org/10.3389/fpsyg.2022.997752

Demos, A. P., & Palmer, C. (2023). Social and nonlinear dynamics unite: Musical group synchrony. *Trends in Cognitive Sciences*, *29*(11), 1008–1018. https://doi.org/10.1016/j.tics.2023.05.005

Endestad, T., Godøy, R. I., Sneve, M. H., Hagen, T., Bochynska, A., & Laeng, B. (2020). Mental effort when playing, listening, and imagining music in one pianist's eyes and brain. *Frontiers in Human Neuroscience*, *14*, 576888. https://doi.org/10.3389/fnhum.2020.576888

Glowinski, D., Bracco, F., Chiorri, C., & Grandjean, D. (2016). Music ensemble as a resilient system: Managing the unexpected through group interaction. *Frontiers in Psychology*, *7*, 1548. https://doi.org/10.3389/fpsyg.2016.01548

Goebl, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception*, *26*(5), 427–438. https://doi.org/10.1525/mp.2009.26.5.427

Hilt, P. M., Badino, L., D'Ausilio, A., Volpe, G., Tokay, S., Fadiga, L., & Camurri, A. (2019). Multi-layer adaptation of group coordination in musical ensembles. *Scientific Reports*, *9*(1), 5854. https://doi.org/10.1038/s41598-019-42395-4

Høffding, S., Yi, W., Lippert, E., Sanchez, V. G., Bishop, L., Laeng, B., Danielsen, A., Jensenius, A. R., & Wallot, S. (2023). Into the hive-hind: Shared absorption and cardiac interrelations in expert and student string quartets. *Music & Science*, *6*, 1–15. https://doi.org/10.1177/20592043231168597

Jakubowski, K., Eerola, T., Blackwood Ximenes, A., Ma, W. K., Clayton, M., & Keller, P. E. (2020). Multimodal perception of interpersonal synchrony: Evidence from global and continuous ratings of improvised musical duo performances. *Psychomusicology: Music, Mind, and Brain*, *30*(4), 159–177. https://doi.org/10.1037/pmu0000264

Kahneman, D. (1973). *Attention and effort*. Prentice-Hall.

Kang, O., & Banaji, M. R. (2020). Pupillometric decoding of high-level musical imagery. *Consciousness and Cognition*, *77*, 102862. https://doi.org/10.1016/j.concog.2019.102862

Kawase, S. (2014). Gazing behavior and coordination during piano duo performance. *Attention, Perception, & Psychophysics*, *76*(2), 527–540. https://doi.org/10.3758/s13414-013-0568-0

Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1658), 20130394. https://doi.org/10.1098/rstb.2013.0394

Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology*, *63*(11), 2220–2230. https://doi.org/10.1080/17470218.2010.497843

Laeng, B., & Alnaes, D. (2019). Pupillometry. In C. Klein & U. Ettinger (Eds.), *Eye movement research* (pp. 449–502). Springer International Publishing.

Laroche, J., Tomassini, A., Volpe, G., Camurri, A., Fadiga, L., & D'Ausilio, A. (2022). Interpersonal sensorimotor communication shapes intrapersonal coordination in a musical ensemble. *Frontiers in Human Neuroscience*, *16*, 899676. https://doi.org/10.3389/fnhum.2022.899676

Lartillot, O., Eerola, T., Toiviainen, P., & Fornari, J. (2008). Multi-feature modeling of pulse clarity: Design, validation and optimization. In *Proceedings of the 9th International Conference on Music Information Retrieval*. Drexel University.

Lehmann, A. C., Gruber, H., & Kopiez, R. (2018). Expertise in music. In K. A. Ericsson, R. R., Hoffman , A. Kozbelt, & A. M. Williams (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (2nd ed., pp. 535–549). Cambridge University Press.

Massie-Laberge, C., Cossette, I., & Wanderley, M. M. (2019). Kinematic analysis of pianists' expressive performances of romantic excerpts: Applications for enhanced pedagogical approaches. *Frontiers in Psychology*, *9*, 2725. https://doi.org/10.3389/fpsyg.2018.02725

Mathôt, S. (2018). Pupillometry: Psychology, physiology, and function. *Journal of Cognition*, *1*(1), 16. https://doi.org/10.5334/joc.18

Moelants, D., Demey, M., Grachten, M., Wu, C.-F., & Leman, M. (2012). The influence of an audience on performers: A comparison between rehearsal and concert using audio, video and movement data. *Journal of New Music Research*, *41*(1), 67–78. https://doi.org/10.1080/09298215.2011.642392

Müller, M. (2021). *Fundamentals of Music Processing: Using Python and Jupyter Notebooks*. Springer International Publishing.

O'Shea, H., & Moran, A. (2018). Are fast complex movements unimaginable? Pupillometric studies of motor imagery in expert piano playing. *Journal of Motor Behavior*, *51*(4), 371–384. https://doi.org/10.1080/00222895.2018.1485010

Palmer, C., Spidle, F., Koopmans, E., & Schubert, P. (2019). Ears, heads, and eyes: When singers synchronise. *Quarterly Journal of Experimental Psychology*, *72*(9), 2272–2287. https://doi.org/10.1177/1747021819833968

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & Team, R. C. (2021). *nlme: Linear and nonlinear mixed effects models*.

Roesch, A., & Schmidbauer, H. (2018). *WaveletComp: Computational Wavelet Analysis. R package version 1.1*.

Schaerlaeken, S., Grandjean, D., & Glowinski, D. (2017). Playing for a virtual audience: The impact of a social factor on gestures, sounds and expressive intents. *Applied Sciences*, *7*(12), 1321. https://doi.org/10.3390/app7121321

Stevens, A., & Ramirez-Lopez, L. (2022). An introduction to the prospectr package. In *R package Vignette*.

Thompson, M. R., & Luck, G. (2012). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, *16*(1), 19–40. https://doi.org/10.1177/1029864911423457

Timmers, R., Endo, S., Bradbury, A., & Wing, A. M. (2014). Synchronization and leadership in string quartet performance: A case study of auditory and visual cues. *Frontiers in Psychology*, *5*, 645. https://doi.org/10.3389/fpsyg.2014.00645

Vandemoortele, S., Feyaerts, K., Reybrouck, M., De Bièvre, G., Brône, G., & De Baets, T. (2018). Gazing at the partner in musical trios: A mobile eye-tracking study. *Journal of Eye Movement Research*, *11*(2), 6. https://doi.org/10.16910/jemr.11.2.6

Weiss, M. W., Trehub, S. E., Schellenberg, E. G., & Habashi, P. (2016). Pupils dilate for vocal or familiar music. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(8), 1061–1065. https://doi.org/10.1037/xhp0000226

Wood, E. A., Chang, A., Bosnyak, D., Klein, L., Baraku, E., Dotov, D., & Trainor, L. J. (2022). Creating a shared musical interpretation: Changes in coordination dynamics while learning unfamiliar music together. *Annals of the New York Academy of Sciences*, *1516*(1), 106–113. https://doi.org/10.1111/nyas.14858

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