



# This is an electronic reprint of the original article. This reprint *may differ* from the original in pagination and typographic detail.

Author(s): Thompson, Marc; Diapoulis, Georgios; Himberg, Tommi; Toiviainen, Petri

Title:Interpersonal Coordination in Dyadic Performance

Year: 2017

Version:

#### Please cite the original version:

Thompson, M., Diapoulis, G., Himberg, T., & Toiviainen, P. (2017). Interpersonal Coordination in Dyadic Performance. In M. Lesaffre, P.-J. Maes, & M. Leman (Eds.), The Routledge Companion to Embodied Music Interaction (pp. 186-194). Routledge, Taylor & Francis Group. Routledge Companions. https://doi.org/10.4324/9781315621364-21

All material supplied via JYX is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

1 **Interpersonal Coordination in Dyadic Performance** 2 3 Marc R. Thompson<sup>1</sup>, Yorgos Diapoulis<sup>1</sup>, Tommi Himberg<sup>2</sup> & Petri Toiviainen<sup>1</sup> 4 Music Department, University of Jyväskylä, Finland 5 <sup>2</sup>Department of Neuroscience and Biomedical Engineering, Aalto University, Finland 6 7 Abstract 8 Dyadic musical performance provides an excellent framework to study interpersonal 9 coordination because it involves multiple agents performing matched, rhythmic and/or interactive behaviors. In this chapter, we explore interpersonal coordination using 10 11 Canonical Correlation Analysis as a coupling measure. To provide some context when 12 interpreting the output of CCA, musicians performed using different expressive manners (deadpan, normal, exaggerated). Overall the results showed the normal 13 14 performances were slightly more interpersonally coordinated than deadpan and 15 exaggerated.

16

#### 17 Introduction

18 Dyadic musical performance provides an excellent framework to study interpersonal 19 coordination because it involves multiple agents performing matched, rhythmic and/or 20 interactive behaviors (Keller, 2008; Himberg & Thompson, 2011). Such behaviors fall under 21 the category of *entrainment*. As Tommi Himberg explains in chapter XX of this book, the 22 criteria of an entrained process are flexibility (adapting one's behavior to match the behavior 23 of their performing partner), autonomy (a performance may continue even if one violinist 24 makes a mistake or stops playing), and coupling (bi-directional communication leading to 25 matched behavior). The advent of motion capture technology has brought with it a growing 26 interest in developing measures that can identify entrained processes computationally. In this 27 chapter we present the results of a small methods-based study in which we explore how 28 Canonical Correlation Analysis (CCA) can be used as a measure of interpersonal 29 coordination within violinist dyads. Because CCA identifies linear relationships between sets 30 of data, it is well suited to measure the extent to which musicians synchronize their behavior 31 when performing. As this is a novel application of CCA, we will examine its results within the context of the Total Kinetic Energy (TKE) employed by the musicians during a
performance. Like CCA, TKE is computed from motion capture data and provides a global
indicator of performance behavior.

35

36 Let's begin with a small thought experiment that we hope will explain the rationale for using 37 CCA. Suppose that you are attending a concert (it could be a classical ensemble, rock show, 38 jazz quartet, anything). Walking out of the concert you are accosted by an intrepid young 39 music researcher and asked to rate on a scale from 1 to 7 the amount of physical interpersonal 40 coordination displayed between the performers. You give a response that from your 41 perspective best sums up the performance. As a follow-up, the researcher wants to know how 42 you arrived to your numerical response and asks you to dissect the coordination in terms of 43 body parts. Were the musicians' heads moving in synchrony? Were the bowing gestures coupled? Perhaps the musicians had similar swaying patterns? Did the musicians 44 45 communicate tempo changes with specific gestures? Were the musicians tapping their feet to 46 keep the beat? Because of the sheer number of events occurring simultaneously, you might 47 feel hard-pressed to describe exactly which parts of the body were chiefly responsible for the 48 perceived coordination. One explanation for this is that while you were not paying attention 49 to individual parts of the body, your mind formed a gestalt of the performance and your 50 rating of interpersonal coordination was based on the performance as a whole, not its 51 constitutional elements.

52

The aim of this thought experiment is to suggest that we are good at recognizing 53 54 interpersonal coordination between musicians even though we are unable to keep track of 55 every piece of information. With technologies such as optical motion capture, we are able to 56 amass vast amounts of data and compare the movement trajectories of markers attached to 57 two musicians in countless ways. We can for instance, create a correlation matrix that 58 compares all markers in all combinations, and observe that Violinist 1's left elbow velocity is 59 highly synchronized with Violinist 2's right knee velocity. However, just as it is more 60 cognitively economical for humans to construct a gestalt that captures the performance as a 61 whole, it is more efficient to compare higher-level movement features that summarize the 62 violinists' movements as a whole. The great advantage of a coupling measure such as CCA is

that we are not comparing marker x with marker y, but comparing Violinist 1 to Violinist 2.
Hence, we can compute the relationship between musicians as a whole with a single
calculation.

66

67 To explain how CCA works, it might be useful to begin with a related technique: Principal Components Analysis (PCA). Consider an  $m \times 3n$  matrix of motion capture data whereby 3n68 69 represents the number of markers attached to a violinist in a three-dimensional coordinate 70 system, and m is the number of observations (in the current case, this number would be the motion capture's sampling rate (120) times the duration of the performance in seconds). 71 72 Collectively, the data set contains fingerings, bowings (i.e. movements involved in sound 73 production) as well as head gestures and torso swaying (i.e. movements that enable sound-74 producing movements). However, much of the data is co-varying (i.e. the head marker's 75 motion is anchored to that of the shoulder markers). PCA is a method that eliminates 76 redundancy embedded within data by creating a new matrix of synthetic orthogonal variables 77 (called the principal components). These variables are linear combinations of the original 78 variables and ordered according to decreasing proportions of variation within the original 79 data matrix (Toiviainen, Luck & Thompson, 2010). The majority of the variation is contained 80 within the first few variables and the remaining variables can be discarded as noise. When 81 applied to motion capture, PCA provides a higher-order movement feature, one that describes 82 the violinist's global dynamic motion rather than the individual markers' motion. As already 83 mentioned, the variables are orthogonal, meaning that they are non-overlapping and 84 uncorrelated. For motion capture data, this means that each principal component will 85 represent the movement occurring on a single axis. For example, if a violinist's predominant 86 movement pattern is side-to-side swaying, the first principal component will include loadings 87 from any marker that has high variation on the mediolateral axis. The extent to which how 88 many additional components should be retained depends on the complexity of the violinist's 89 movements.

90

91 CCA is similar to PCA but takes into account *two* data matrices, and computes their *shared*92 variance. In our case, the matrices represent the movements of each violinist. As such, CCA
93 produces *two* synthetic data sets. These synthetic variables (called the canonical components)

94 reflect the variance that is shared with the other matrix. The canonical components are 95 ordered according to decreasing amounts of shared variance; the first canonical components 96 from each data set contain the highest amount of shared variance to maximize their 97 correlation. Like PCA, only the first few canonical components should be interpreted as 98 meaningful, and the remaining can be discarded as noise. In relation to the current study, 99 CCA offers two primary advantages (Sherry & Henson, 2005). First, as a multivariate 100 technique, it greatly diminishes the probability of committing Type I errors. Correlating 101 everything with everything is likely to produce false positives-CCA decreases the chances 102 of this happening because it is based on a single correlation. Second, CCA enables one to 103 observe how multiple variables interact with each other. Human behavior is complex and governed by multiple causes and effects. While univariate methods test each variable in 104 105 isolation (i.e. how marker x is related to marker y), CCA offers the potential to examine 106 multivariate relationships (i.e. how Violinist 1 is related to Violinist 2). CCA has previously 107 been used in studies using motion capture and music to investigate relationships between 108 sound-tracing gestures and computationally extracted acoustical features (Caramiaux, 109 Bevilacqua & Schnell, 2010; Nymoen et al, 2013). In contrast, we are using CCA to identify 110 linear relationships between the gestures of performing musicians, and have not taken into 111 account any acoustical features.

112

113 In this chapter, we also take into account the coordination of the violinists' affective states 114 and how this might affect the results of the CCA. Research targeting the relationship between 115 affective states and body movement has come of age in the past 20 years, particularly in 116 studies on solo performance. Starting with Davidson's seminal 1993 study, a popular design 117 has been to instruct musicians to perform with varying expressive manners (Davidson, 1993; 118 Palmer et al., 2009; Wanderley et al., 2005; Huang & Krumhansl, 2011; Thompson & Luck, 119 2012)-manners might include: deadpan (i.e., without expression), projected (i.e., with 120 normal levels of expression), and exaggerated (i.e., with exaggerated levels of expression). 121 Davidson's (1993) study focused on the perception of expressive manner revealed that less 122 experienced observers were more likely to use the visual modality when rating level of 123 musical expressivity. More recent perceptual studies have further investigated the crossmodal 124 interactions of vision and sound that observers employ when making expressivity judgments

125 (Vuoskoski et al., 2013). The 'levels of expression paradigm' has also been used in music 126 production experiments in which the goal has been to analyze the kinematics of pianists 127 whose movements have been motion captured. Thompson and Luck's (2012) data showed 128 that the head and shoulders had bigger differences between expressive manners, compared to the fingers and wrists. The authors argued that pianists might use the parts of the body not 129 130 involved in sound production as a means of conveying expressivity, as these would have 131 greater degrees of freedom than parts of the body directly involved in sound production. In 132 recent years, an increased accessibility of sensor technologies, as well as an advancement 133 towards studying music performance from the lens of embodied cognition (Leman, 2007), 134 has prompted a number of studies on interpersonal coordination within musical dyads to use 135 motion capture for data collection (Goebl & Palmer, 2009; Keller & Appel, 2010; also see 136 Repp & Su, 2013).

137

#### 138 The Current Study

139 Our aim for this chapter is to contribute to work on interpersonal coordination in dyadic 140 performance by using canonical correlation analysis (CCA) for quantifying coupled 141 movements in a musical performance. Below we report a small study in which three violinist 142 dyads were motion-captured while performing a short piece. Our focus is on non-sound-143 producing gestures (sometimes called ancillary gestures; Wanderley et al., 2005). Ancillary 144 gestures are best represented through low-level kinematic features such as velocity. For a 145 coupling measure like CCA, velocity is more appropriate than position for two main reasons. First, it eliminates the position factor. For instance, if one waves their hand at two different 146 147 locations, the velocity data are similar while the position data are dissimilar. Second, velocity is mostly stationary (e.g. zero-centered), whereas position may not be. Nonstationarity in 148 149 position data (e.g. a significant change of posture) may be the single determining factor in 150 correlation, which may hide other, possibly important but more nuanced forms of coupling. 151 Because CCA is novel in this context, we compare its results with Total Kinetic Energy, 152 which has been used previously to summarize performance behavior (Toiviainen et al., 153 2010). As to the experimental design, we felt it would be useful to examine how the results 154 of canonical correlation analysis might vary under different performance conditions. Because 155 our previous work has shown that musicians alter their movement patterns when requested to perform in different levels of expression (Thompson & Luck, 2012), and in different tempi (Thompson et al., 2015), we hypothesized that interpersonal coordination would also be different according to instructed expressive intention and tempo, and that these differences would be reflected in the results of canonical correlation analysis.

160

## 161 Method

# 162 Participants

163 Three violin dyads participated in this study (6 musicians total; 4 females; age: M = 24.1, SD164 = 1.7). The violinists were recruited from student populations at the University of Jyväskylä 165 and JAMK University of Applied Sciences. Musicians had received on average 15.8 (SD =166 2.3) years of instrumental training on the violin. Although the violinists knew each other, 167 none had spent any significant amounts of time performing together before to the experiment. 168

169 Procedure

The dyads performed a short piece arranged for two violins: *De Kleinste*, composed by J. Beltjens (16 bars, 6/8 time signature). Each dyad performed the piece nine times in a 3×3 task design: three expressive manners (*deadpan, normal, exaggerated*) performed using three tempi (60-BPM, 90-BPM, free tempo). Our data set thus consists of 27 performances ( $M_{duration} =$ 33.1 sec,  $SD_{duration} = 4.6$  sec).

175

176 The tempo was given using a two-measure metronome count-in. Despite instructing a tempo, the dyads gravitated to their own tempo, abandoning the instructed tempo within a few bars. 177 178 To explore the deviation from the instructed tempo, we estimated the actual tempo for each 179 performance. Due to timing dynamics, the tempi also fluctuated within the performances; to 180 calculate a performance's average tempo, we divided the performance's duration in seconds 181 by the number of eighth note values in the score. We then converted this time in seconds per 182 meter to BPM. Generally, the dyads performed faster than the instructed tempo (actual 60-183 BPM: M = 76.56 BPM, SD = 5.66; actual 90-BPM: M = 94.78 BPM, SD = 12.35). For the 184 Free Tempo condition, the average tempo was 77.1 BPM (SD = 7.16). Because the Free Tempo was so close to the actual tempo for 60-BPM, we estimate that the overall preferred 185 186 performing tempo was about 76 or 77 BPM (the tempo marking on the score for De Kleinste. 187 50 BPM for each dotted quarter note). Initial analysis exploration indicated that tempo did 188 not interact in any meaningful way with the expressive manners. Therefore, we eliminated 189 tempo as a factor and all of the results presented below represent data that is averaged by the 190 expressive manner.

191

## 192 Audio Recordings

Audio of the experimental trials was recorded using two AKG C417 L wireless microphones. The microphones were positioned around each violinist's right ear lobe and secured with adhesive tape. Note onsets were annotated from the recorded audio files in Reaper digital audio workstation manually using the marker utilities extension. Each take was exported as a plain text file with note onsets in samples.

198

199 *Motion Capture* 

200 Optical motion capture data was generated using eight Qualisys Oqus cameras at 120 Hz. 201 Twenty-six markers were placed on the joints of each musician, and five markers were placed 202 on the violin (2 on the bow, and 3 on the violin itself). The data were labeled within 203 Qualisys' Track Manager software and analyzed in MATLAB using functions within the 204 MoCap Toolbox (Burger & Toiviainen, 2013). CCA was carried out using MATLAB's 205 *canoncorr* function.

206

#### 207 Feature Computation

#### 208 Total Kinetic Energy

209 For every performance we computed the instantaneous kinetic energy for each performer 210 individually based on the method used by Toiviainen et al. (2010). The total kinetic energy 211 was estimated as the sum of the kinetic energy of both performers, which consists of the total 212 amount of translational and rotational energy computed for each marker. Specifically, we 213 used the function mckinenergy from MoCap Toolbox (Burger & Toiviainen, 2013) to 214 compute both the translational and rotational energy based on a body segment model proposed by Dempster (1959). We then calculated the mean kinetic energy for each marker, 215 216 and took the grand sum for both performers. Total kinetic energy (TKE) describes the amount 217 of physical activity across both performers in one performance.

218

#### 219 Canonical Correlation Analysis: CanCon1 and CanCon2

220 CCA was employed on each performance separately. Each performance produced two 221 matrices (V1 and V2) representing the three-dimensional coordinates of 28 reflective markers 222 attached to each violinist. Because we were interested in coupled ancillary movement (e.g. 223 swaying), we first reduced the data to seven markers representing non-sound producing 224 movement patterns of each violinist: head, left shoulder, right shoulder, neck, left hip, right 225 hip and hip centroid. From the raw three-dimensional location data, we estimated the velocity 226 of V1 and V2 using numerical differentiation. To this end, we applied finite difference 227 followed by a second-order smoothing Butterworth filter with a 0.2 Hz cutoff frequency (see 228 Burger & Toiviainen, 2013). As mentioned above, the markers attached to a musician 229 generate data that are largely co-varying. Entering the velocity data from all seven markers 230 into the CCA equates to including noise to the CCA, and results in an over-fitted model. For 231 this reason, Principal Components Analysis (PCA) was applied to V1 and V2 to reduce the 232 number of input variables. On average, we found the first two principal components retained 233 85.5% of the total variance within the raw movement data ( $M_{\rm rc1} = 72.6\%$ ,  $SD_{\rm rc1} = 14.27\%$ ;  $M_{\rm rc2}$ = 13.26%,  $SD_{\text{res}} = 7\%$ ,). The transformed data, called the PC scores (original data expressed 234 235 as values of the reduced set of variables), were entered in the CCA. We name these input 236 variables PCs1 and PCs2.

237

238 The outputs of CCA are the canonical loadings (A and B), which are the sets of coefficients 239 that indicate the contributions of the variables from PCs1 and PCs2 to the new sets of 240 variables, and the *canonical scores* (U and V), which are the new variables themselves. Like 241 in PCA, the canonical scores are synthetic data sets in which the column variables are 242 orthogonal to each other. Instead of being ordered according to variance within a single set of 243 variables, the canonical scores are ordered according to the variance shared between PCs1 244 and PCs2. Therefore the highest correlation exists between the respective first canonical 245 components (CanCon1) as they contain as much of the shared variance as possible between 246 PCs1 and PCs2. The second canonical components (CanCon2) in turn account for as much of the remaining shared variance as possible. The strength of the relationship between the 247 248 canonical components is represented by the sample correlation between respective columns

of U and V. In CCA, the correlation values range from 0 to 1. A value closer to 1 implies that violinists coupled their movements, or were to some extent synchronized.

It is critical to mention that when using CCA for timeseries movement data, we must take into account that the data are autocorrelated, and do not meet the assumption of serial independence between observations required for parametric tests. Specifically, the sample correlation between two variables containing autocorrelated observations has fewer degrees of freedom than two variables containing serially independent observations. In other words, in a timeseries, the number of observations that are free to vary is presumably less than N - 1, where N is the number of observations. To overcome this obstacle, we estimated the *effective* degrees of freedom based on the autocorrelation functions of U and V, using a formula that takes into account the autocorrelation coefficient appearing at the maximum lag (Pyper and Peterman, 1998; Alluri et al., 2012). Averaged for all 27 performances, we calculated the effective degrees of freedom for CanCom1 ( $M_{cc1} = 47.28$ ,  $SD_{cc1} = 27.42$ ) and CanCom2 ( $M_{cc2} =$ 116.96,  $SD_{cc} = 187.23$ ), which were in turn used to compute the correlation *p*-values. For CanCon1, all performances were significant (p < .05), whereas for CanCom2, none of the performances were significant. 



FIGURE X: Summary (mean and standard deviation) of for A) Total Kinetic Energy (*TKE*), B)
CanCom1 (*r*), C) CanCom2 (*r*). Panels D, E and F plot canonical scores over time.

283

#### 284 Feature Comparison

285 Total Kinetic Energy: Figure X, panel A shows the mean TKE for deadpan, normal and 286 exaggerated performances, with error bars to represent standard deviation. The results show 287 that different expressive manners are reflected in the total amount of movement, with greater 288 amount of movement observed as expression level increases (deadpan: lowest, exaggerated: 289 highest). These results can be viewed as a proof of concept for the expressive manners: 290 Musicians altered their behavior when asked to perform with more expression. It should be 291 noted that the musicians were never overtly instructed to change their movement patterns 292 according to the expressive manner. Rather, the different behaviors are natural emergent 293 consequences of the expressive manners.

294

295 CanCom1 & CanCom2: Panels B and C show the mean sample correlations for deadpan, 296 normal and exaggerated for CanCom1 and CanCom2 respectively. Panel B shows that each 297 expressive manner was performed with roughly similar amounts of coupled movements. 298 However, normal performances are slightly more coupled than the other two. One 299 interpretation is that normal performances did not require extra cognitive load on the part of 300 the players. That is, they could perform without taking into consideration what it meant to 301 play deadpan or exaggerated.

302

303 Panels B and C also indicate information regarding movement complexity. It appears that the 304 shared variance within deadpan and normal performances is essentially entirely explained by 305 CanCom1. This could be interpreted by stating on average, interpersonal coupling in these 306 performances was occurring on a single axis. In other words, the movement associated with 307 coupling was not complex. For the exaggerated performances, the gap is smaller between 308 CanCom1 and CanCom2, meaning that the shared variation is distributed across the two 309 canonical components. Taking into account that exaggerated performances possessed the 310 most kinetic energy (panel A), we can make the assertion that the interpersonal coordination 311 in the exaggerated performances was generally more complex, and occurring on multiple 312 axes.

#### 314 Exploring the Canonical Loadings and Canonical Scores

Now let us examine more closely the canonical loadings and canonical scores. For this section, we have singled out performances that best exemplify each expressive manner. We recommend the reader to watch the performances in this chapter's accompanying video (found at <u>https://www.youtube.com/watch?v=fvv2-he5lco&t=67s</u>).

319

320 Given that we have performed canonical correlation analysis on the first two principal 321 components of seven velocity features that were estimated from three-dimensional marker 322 location data (gasp!), you may be asking the very sensible question: After so much 323 processing, how do the canonical scores and loadings relate to the original motion capture 324 data? To answer this question, let us first consider what information the loadings and scores 325 provide. Since we are dealing with motion capture data, an intuitive way to think about the 326 loadings and the scores is that the scores provide temporal information while the loadings 327 provide spatial information.

328

329 Canonical scores: The canonical scores represent temporal information because they are 330 timeseries with the same amount of observations as the original location data. If we overlay 331 the canonical scores from each violinist, we get an impression as to how the coupling evolves 332 throughout the performance. In panels D, E and F, we have plotted the CanCom1 canonical 333 scores from performances that best represent low and high gesture coupling. Panel D shows 334 the canonical score from the deadpan performance with the lowest CanCom1 value (r = .23) 335 out of the 27 performances (first lowest overall). Panel E shows the canonical score from the 336 exaggerated performance with the lowest CanCom1 value (r = .44) out of the 27 337 performances (sixth lowest overall). Panel F shows the normal performance with the highest 338 CanCom1 value (r = .87) and highest over all 27 performances. Comparing the three panels, 339 we see that the deadpan performance scores are irregular and do not co-vary coupling in the 340 normal performance (panel F) is regular with some deviations after the 25-second mark. 341 Meanwhile, the

342

343 *Canonical loadings*: Recall that the loadings indicate the various contributions of the input 344 variables (PCs1 and PCs2) to the output variables (U and V). These contributions can be 345 inspected visually through the use of animations. In our accompanying video, we have 346 created animations that represent the loadings of CanCom1 and CanCom2 for each example 347 performance. The animations were created by projecting the canonical loadings back to the 348 kinematic (velocity) space by multiplying the PC loading matrix with the loading matrix of 349 the two canonical components and adding the mean position back to the data. The animations should not be seen as being time-dependent; they demonstrate the extent to which the various 350 351 markers contribute to the canonical components. The extent to which a marker is moving 352 indicates its contribution to the canonical component. The animations also demonstrate how canonical components are orthogonal, and so the variability within a canonical component is 353 354 largely restricted to a single axis.

355 356

#### 357 Conclusions

358 At the beginning of this chapter, we mentioned that an entrained process should contain 359 flexibility, autonomy and coupling. In truth, employing CCA the way we did does not allow 360 us to identify entrained processes because we cannot account for flexibility and autonomy. 361 While CCA provides a measure of linear coupling, it only gives a global measure, ignoring 362 the time-dependent aspects of entrained processes. Our earlier studies on entrainment in 363 dance performance within dyads (Himberg & Thompson, 2010; Himberg & Thompson, 2011) found that expert/novice dyads remained synchronized not by appropriating a 364 365 leader/follower dynamic, but through mutual adaptation in response to each other's 366 movements. This does not rule out that entrainment never occurred within our data set. On 367 the contrary, looking at panels D and E, we can identify some epochs within the timeseries 368 that appear more together than others. This variation within the canonical scores indicates 369 that performers adapted their behaviors throughout the performance. Therefore, CCA might 370 have the potential to identify entrained processes if for example, the scores were windowed 371 and multiple comparisons were performed laterally on a single time series. Additionally, 372 cross-correlation analysis could identify the direction of the interaction and reveal 373 leader/follower relationships. Using these alterations to our method, we may be able to use 374 CCA more effectively for identifying dynamic entrainment processes.

376 There are significant drawbacks to CCA for motion capture data. As explained above, the 377 serial correlation within the data requires the use of effective degrees of freedom for 378 calculating *p*-values. The greater amount of serial correlation leads to a lower amount of 379 degrees of freedom (the performance in panel F contains only 11 effective degrees of freedom for over 2000 observations). Another drawback is that CCA can only identify linear 380 381 relationships between two data sets; it is ineffective for detecting processes that are non-382 linear. For detecting non-linear relationships, there exists a broad range of coupling 383 techniques other than CCA. For example, mutual information is a technique that identifies the 384 mutual dependence between two data sets. While mutual information would do a better job of 385 finding non-linear relationships between sets, it is a bin-based approach that requires 386 knowing how many bins to use before the calculation. As such, the result is greatly affected 387 by input parameters. In this respect CCA is advantageous because it is parameter-free and 388 simpler to compute.

389

390 To summarize, we have examined Canonical Correlation Analysis as a means of quantifying 391 coupled movement in performing dyads. To provide a context when interpreting the output of 392 CCA, musicians performed using different expressive manners (deadpan, normal, 393 exaggerated). We also calculated the total kinetic energy across the expressive manners, 394 which provided evidence that musicians altered their behavior based on the performance 395 instruction. Overall the results showed the normal performances were slightly more 396 interpersonally coordinated than deadpan and exaggerated. While we cannot claim that our 397 exact results would be replicated (music performance is an idiosyncratic endeavor), we hope 398 we have presented CCA in sufficient detail to inspire others interested in interpersonal 399 coordination to explore this coupling measure.

- 400
- 401

#### 402 **References**

- 403 Alluri, V., Toiviainen, P., Jääskeläinen, I. P., Glerean, E., Sams, M., & Brattico, E. (2012).
- Large-scale brain networks emerge from dynamic processing of musical timbre, key and rhythm. *Neuroimage*, *59*, 677–3689.
- Burger, B. & Toiviainen, P. (2013). Mocap toolbox a matlab toolbox for computational
  analysis of movement data. In Bresin, R., (Ed.), *Proceedings of the 10th Sound and Music*
- 408 Computing Conference (pp. 172–178), Stockholm, Sweden. KTH Royal Institute of
   409 Technology.
- 410 Caramiaux, B., Bevilacqua, F., & Schnell, N. (2010). Towards a gesture-sound cross-modal
  411 analysis. In *Gesture in Embodied Communication and Human-Computer Interaction*, (pp.
  412 158–170). Springer.
- 413 Davidson, J. (1993). Visual perception of performance manner in the movements of solo
  414 musicians. *Psychology of Music*, 21, 103–113.
- Dempster, W. T., Gabel, W. C., & Felts, W. J. L. (1959). The anthropometry of manual work
  space for the seated subject. *American Journal of Physical Anthropology*, *17*, 289–317.
- Goebl, W. & Palmer, C. (2009). Synchronization of timing and motion among performing
  musicians. *Music Perception*, 26, 427–425.
- 419 Himberg, T. & Thompson, M. (2010). Dyadic entrainment and interaction in african dance.
- 420 In Demorest, S., Morrison, S., and Campbell, P., (Eds.) Proceedings of the 11th
- 421 international conference for music perception and cognition, (ICMPC '11), (pp. 425–
- 422 428), Seattle, USA. University of Washington.
- Himberg, T. & Thompson, M. R. (2011). Learning and synchronizing dance movements in
  South African songs: Cross-cultural motion-capture study. *Dance Research*, 29(special *electronic issue*), 305–328.
- 426 Hotelling, H. (1936). Relations between two sets of variates. *Biometrika*, 28, 321–377.
- Huang, J. and Krumhansl, C. L. (2011). What does seeing the performer add? It depends on
  musical style, amount of stage behavior, and audience expertise. *Musicæ Scientiæ*, 15,
  343–364.
- 430 Keller, P. E. (2008). Joint action in music performance. In Morganti, F., Carassa, A., and
- 431 Riva, G., (Eds.) Enacting Intersubjectivity: A Cognitive and Social Perspective on the
- 432 *Study of Interactions*, (pp. 205–221). IOS Press, Amsterdam.

- Keller, P. E. (2013). Ensemble performance: Interpersonal alignment of musical expression.
  In Fabian, D., Timmers, R., and Schubert, E., (Eds.), *Expressiveness in music performance: Empirical approaches across styles and cultures*. Oxford University Press.
- Keller, P. E. & Appel, M. (2010). Individual differences, auditory imagery, and the
  coordination of body movements and sounds in musical ensembles. *Music Perception*, 28,
  27–46.
- 439 Leman, M. (2008). *Embodied music cognition and mediation technology*. Cambridge, MA:
  440 MIT Press.
- 441 Nymoen, K., Godøy, R.I., Jensenius, A.R., & Torresen, J. (2013). Analyzing correspondence
  442 between sound objects and body motion. *ACM Transactions on Applied Perception*, *10*,
  443 article 9, 1-22.
- 444 Palmer, C., Koopmans, E., Carter, C., Loehr, J., & Wanderley, M. (2009). Synchronization of
- 445 motion and timing in clarinet performance. In Williamon, A., Pretty, S., and Buck, R.,
- 446 (Eds.), Proceedings of the International Symposium on Performance Science, (pp. 159–
- 447 164). European Association of Conservatories (AEC), Utrecht.
- 448 Pyper, B. J. & Peterman, R. M. (1998). Comparison of methods to account for
  449 autocorrelation in correlation analyses of fish data. *Canadian Journal of Fisheries and*450 *Aquatic Sciences*, 55, 2127–2140.
- 451 Repp, B. H. & Su, Y.-H. (2012). Sensorimotor synchronization: A review of recent research
  452 (2006-2012). *Psychonomic Bulletin & Review*, 20, 403-452.
- 453 Sherry, A. & Henson, R. K. (2005). Conducting and interpreting canonical correlation
  454 analysis in personality research: A user-friendly primer. *Journal of Personality*455 Assessment, 84, 37–48.
- Thompson, M. R., Diapoulis, G., Johnson, S., Kwan, P. Y., & Himberg, T. (2015). Effect of
  tempo and vision on interpersonal coordination of timing in dyadic performance. In
  Aramaki, M., Kronland-Martinet, R., and Ystad, S., (Eds.), *Proceedings of the 11th International Symposium on Computer Music Multidisciplinary Research*. University of
  Plymouth, UK.
- Thompson, M. R. & Luck, G. (2012). Exploring relationships between pianists' body
  movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, *16*, 19–40.

- 464 Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical
  465 eigenmodes in music-induced movement. *Music Perception*, 28, 59–70.
- Vuoskoski, J., Thompson, M., Spence, C., & Clarke, E. (2013). Crossmodal interactions in
  the perception of expressivity in musical performance. *Attention, Perception, &*
- 468 *Psychophysics*, 76, 591–604.
- 469 Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C., & Hatch, W. (2005). The
- 470 musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal*471 *of New Music Research*, *34*, 97–113.
- 4/1 0j New Music Research, 54, 97–115.