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

Interpersonal Coordination in Dyadic Performance

Marc R. Thompson[1], Yorgos Diapoulis[1], Tommi Himberg[2] & Petri Toiviainen[1]
[1]Music Department, University of Jyväskylä, Finland
[2]Department of Neuroscience and Biomedical Engineering, Aalto University, Finland

Abstract

Dyadic musical performance provides an excellent framework to study interpersonal coordination because it involves multiple agents performing matched, rhythmic and/or interactive behaviors. In this chapter, we explore interpersonal coordination using Canonical Correlation Analysis as a coupling measure. To provide some context when interpreting the output of CCA, musicians performed using different expressive manners (deadpan, normal, exaggerated). Overall the results showed the normal performances were slightly more interpersonally coordinated than deadpan and exaggerated.

Introduction

Dyadic musical performance provides an excellent framework to study interpersonal coordination because it involves multiple agents performing matched, rhythmic and/or interactive behaviors (Keller, 2008; Himberg & Thompson, 2011). Such behaviors fall under the category of *entrainment*. As Tommi Himberg explains in chapter XX of this book, the criteria of an entrained process are flexibility (adapting one's behavior to match the behavior of their performing partner), autonomy (a performance may continue even if one violinist makes a mistake or stops playing), and coupling (bi-directional communication leading to matched behavior). The advent of motion capture technology has brought with it a growing interest in developing measures that can identify entrained processes computationally. In this chapter we present the results of a small methods-based study in which we explore how Canonical Correlation Analysis (CCA) can be used as a measure of interpersonal coordination within violinist dyads. Because CCA identifies linear relationships between sets of data, it is well suited to measure the extent to which musicians synchronize their behavior 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.

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

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

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 $3n$ represents the number of markers attached to a violinist in a three-dimensional coordinate 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). Collectively, the data set contains fingerings, bowings (i.e. movements involved in sound production) as well as head gestures and torso swaying (i.e. movements that enable sound-producing movements). However, much of the data is co-varying (i.e. the head marker's motion is anchored to that of the shoulder markers). PCA is a method that eliminates redundancy embedded within data by creating a new matrix of synthetic orthogonal variables (called the principal components). These variables are linear combinations of the original variables and ordered according to decreasing proportions of variation within the original data matrix (Toiviainen, Luck & Thompson, 2010). The majority of the variation is contained within the first few variables and the remaining variables can be discarded as noise. When applied to motion capture, PCA provides a higher-order movement feature, one that describes the violinist's global dynamic motion rather than the individual markers' motion. As already mentioned, the variables are orthogonal, meaning that they are non-overlapping and uncorrelated. For motion capture data, this means that each principal component will represent the movement occurring on a single axis. For example, if a violinist's predominant movement pattern is side-to-side swaying, the first principal component will include loadings from any marker that has high variation on the mediolateral axis. The extent to which how many additional components should be retained depends on the complexity of the violinist's movements.

CCA is similar to PCA but takes into account *two* data matrices, and computes their *shared* variance. In our case, the matrices represent the movements of each violinist. As such, CCA produces *two* synthetic data sets. These synthetic variables (called the canonical components) reflect the variance that is shared with the other matrix. The canonical components are ordered according to decreasing amounts of shared variance; the first canonical components from each data set contain the highest amount of shared variance to maximize their correlation. Like PCA, only the first few canonical components should be interpreted as meaningful, and the remaining can be discarded as noise. In relation to the current study, CCA offers two primary advantages (Sherry & Henson, 2005). First, as a multivariate technique, it greatly diminishes the probability of committing Type I errors. Correlating everything with everything is likely to produce false positives—CCA decreases the chances of this happening because it is based on a single correlation. Second, CCA enables one to 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 isolation (i.e. how marker x is related to marker y), CCA offers the potential to examine multivariate relationships (i.e. how Violinist 1 is related to Violinist 2). CCA has previously been used in studies using motion capture and music to investigate relationships between *sound-tracing* gestures and computationally extracted acoustical features (Caramiaux, Bevilacqua & Schnell, 2010; Nymoen et al, 2013). In contrast, we are using CCA to identify linear relationships between the gestures of performing musicians, and have not taken into account any acoustical features.

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

(Vuoskoski et al., 2013). The ‘levels of expression paradigm’ has also been used in music production experiments in which the goal has been to analyze the kinematics of pianists whose movements have been motion captured. Thompson and Luck’s (2012) data showed 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 involved in sound production as a means of conveying expressivity, as these would have greater degrees of freedom than parts of the body directly involved in sound production. In recent years, an increased accessibility of sensor technologies, as well as an advancement towards studying music performance from the lens of embodied cognition (Leman, 2007), has prompted a number of studies on interpersonal coordination within musical dyads to use motion capture for data collection (Goebl & Palmer, 2009; Keller & Appel, 2010; also see Repp & Su, 2013).

The Current Study

Our aim for this chapter is to contribute to work on interpersonal coordination in dyadic performance by using canonical correlation analysis (CCA) for quantifying coupled movements in a musical performance. Below we report a small study in which three violinist dyads were motion-captured while performing a short piece. Our focus is on non-sound-producing gestures (sometimes called ancillary gestures; Wanderley et al., 2005). Ancillary gestures are best represented through low-level kinematic features such as velocity. For a 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 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 position data (e.g. a significant change of posture) may be the single determining factor in correlation, which may hide other, possibly important but more nuanced forms of coupling. Because CCA is novel in this context, we compare its results with Total Kinetic Energy, which has been used previously to summarize performance behavior (Toiviainen et al., 2010). As to the experimental design, we felt it would be useful to examine how the results of canonical correlation analysis might vary under different performance conditions. Because 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.

Method

Participants

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

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).

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. To explore the deviation from the instructed tempo, we estimated the actual tempo for each performance. Due to timing dynamics, the tempi also fluctuated within the performances; to calculate a performance's average tempo, we divided the performance's duration in seconds by the number of eighth note values in the score. We then converted this time in seconds per meter to BPM. Generally, the dyads performed faster than the instructed tempo (actual 60-BPM: M = 76.56 BPM, SD = 5.66; actual 90-BPM: M = 94.78 BPM, SD = 12.35). For the 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 performing tempo was about 76 or 77 BPM (the tempo marking on the score for *De Kleinste*.

50 BPM for each dotted quarter note). Initial analysis exploration indicated that tempo did not interact in any meaningful way with the expressive manners. Therefore, we eliminated tempo as a factor and all of the results presented below represent data that is averaged by the expressive manner.

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.

Motion Capture

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

Feature Computation

Total Kinetic Energy

For every performance we computed the instantaneous kinetic energy for each performer individually based on the method used by Toiviainen et al. (2010). The total kinetic energy was estimated as the sum of the kinetic energy of both performers, which consists of the total amount of translational and rotational energy computed for each marker. Specifically, we used the function *mckinenergy* from MoCap Toolbox (Burger & Toiviainen, 2013) to 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, and took the grand sum for both performers. Total kinetic energy (TKE) describes the amount of physical activity across both performers in one performance.

Canonical Correlation Analysis: CanCon1 and CanCon2

CCA was employed on each performance separately. Each performance produced two matrices (V1 and V2) representing the three-dimensional coordinates of 28 reflective markers attached to each violinist. Because we were interested in coupled ancillary movement (e.g. swaying), we first reduced the data to seven markers representing non-sound producing movement patterns of each violinist: head, left shoulder, right shoulder, neck, left hip, right hip and hip centroid. From the raw three-dimensional location data, we estimated the velocity of V1 and V2 using numerical differentiation. To this end, we applied finite difference followed by a second-order smoothing Butterworth filter with a 0.2 Hz cutoff frequency (see Burger & Toiviainen, 2013). As mentioned above, the markers attached to a musician generate data that are largely co-varying. Entering the velocity data from all seven markers into the CCA equates to including noise to the CCA, and results in an over-fitted model. For this reason, Principal Components Analysis (PCA) was applied to V1 and V2 to reduce the number of input variables. On average, we found the first two principal components retained 85.5% of the total variance within the raw movement data (M_{PC1} = 72.6%, SD_{PC1} = 14.27%; M_{PC2} = 13.26%, SD_{PC2} = 7%,). The transformed data, called the PC scores (original data expressed as values of the reduced set of variables), were entered in the CCA. We name these input variables PCs1 and PCs2.

The outputs of CCA are the *canonical loadings* (A and B), which are the sets of coefficients that indicate the contributions of the variables from PCs1 and PCs2 to the new sets of variables, and the *canonical scores* (U and V), which are the new variables themselves. Like in PCA, the canonical scores are synthetic data sets in which the column variables are orthogonal to each other. Instead of being ordered according to variance within a single set of variables, the canonical scores are ordered according to the variance *shared* between PCs1 and PCs2. Therefore the highest correlation exists between the respective first canonical components (CanCon1) as they contain as much of the shared variance as possible between 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 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 ($M_{cc2} = 116.96$, $SD_{cc2} = 187.23$), 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.

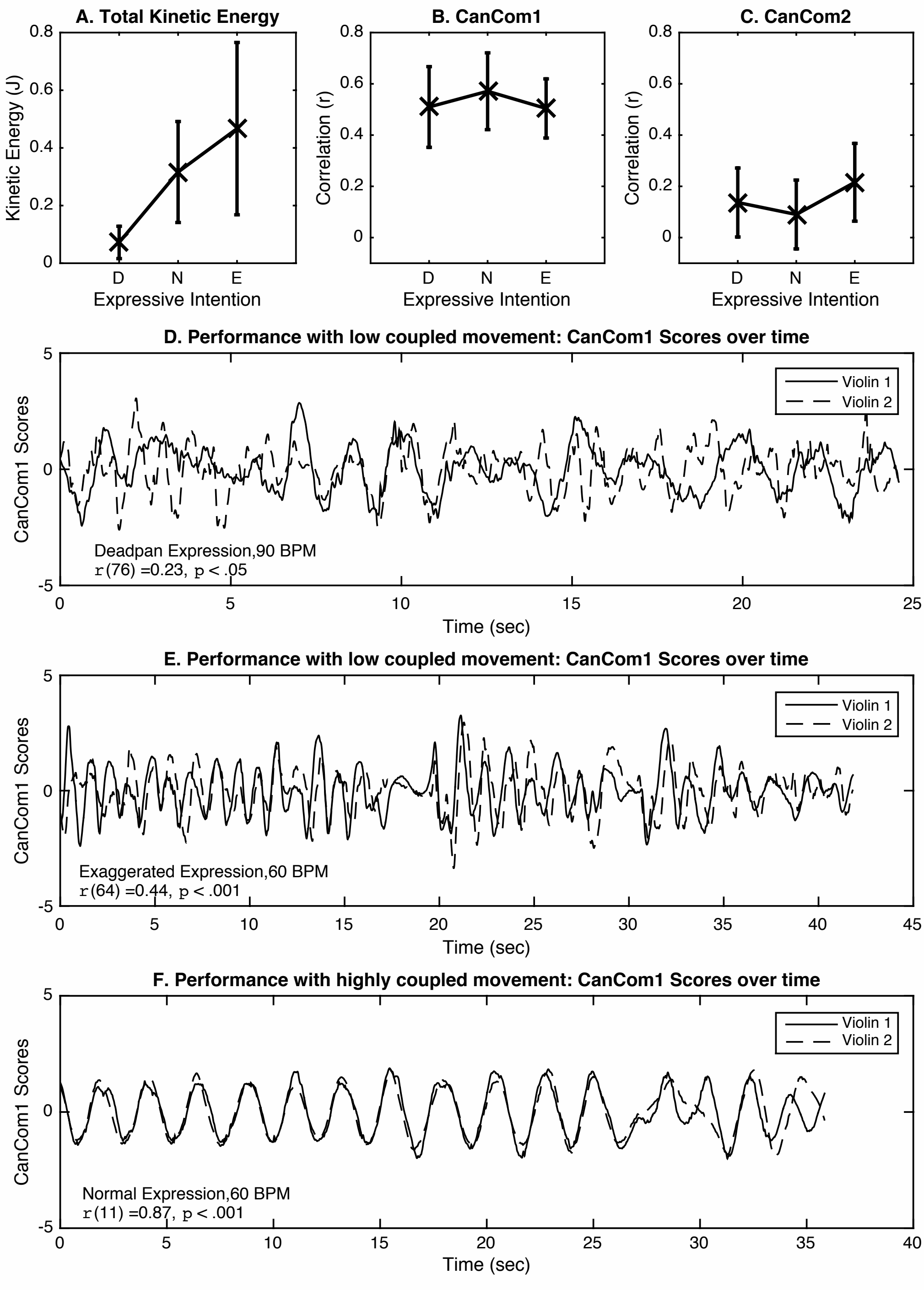

FIGURE X: Summary (mean and standard deviation) of for A) Total Kinetic Energy (*TKE*), B) CanCom1 (*r*), C) CanCom2 (*r*). Panels D, E and F plot canonical scores over time.

Feature Comparison

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

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

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

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 https://www.youtube.com/watch?v=fvv2-he5lco&t=67s).

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

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

Canonical loadings: Recall that the loadings indicate the various contributions of the input variables (PCs1 and PCs2) to the output variables (U and V). These contributions can be inspected visually through the use of animations. In our accompanying video, we have created animations that represent the loadings of CanCom1 and CanCom2 for each example performance. The animations were created by projecting the canonical loadings back to the kinematic (velocity) space by multiplying the PC loading matrix with the loading matrix of 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 markers contribute to the canonical components. The extent to which a marker is moving 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 largely restricted to a single axis.

Conclusions

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

There are significant drawbacks to CCA for motion capture data. As explained above, the serial correlation within the data requires the use of effective degrees of freedom for calculating *p*-values. The greater amount of serial correlation leads to a lower amount of 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 relationships between two data sets; it is ineffective for detecting processes that are non-linear. For detecting non-linear relationships, there exists a broad range of coupling techniques other than CCA. For example, mutual information is a technique that identifies the mutual dependence between two data sets. While mutual information would do a better job of finding non-linear relationships between sets, it is a bin-based approach that requires knowing how many bins to use before the calculation. As such, the result is greatly affected by input parameters. In this respect CCA is advantageous because it is parameter-free and simpler to compute.

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

References

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 Computing Conference* (pp. 172–178), Stockholm, Sweden. KTH Royal Institute of Technology.

Caramiaux, B., Bevilacqua, F., & Schnell, N. (2010). Towards a gesture-sound cross-modal analysis. In *Gesture in Embodied Communication and Human-Computer Interaction*, (pp. 158–170). Springer.

Davidson, J. (1993). Visual perception of performance manner in the movements of solo 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.

Himberg, T. & Thompson, M. (2010). Dyadic entrainment and interaction in african dance. In Demorest, S., Morrison, S., and Campbell, P., (Eds.) *Proceedings of the 11th international conference for music perception and cognition*, (ICMPC ’11), (pp. 425–428), 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.

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.

Keller, P. E. (2008). Joint action in music performance. In Morganti, F., Carassa, A., and Riva, G., (Eds.) *Enacting Intersubjectivity: A Cognitive and Social Perspective on the 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.

Leman, M. (2008). *Embodied music cognition and mediation technology*. Cambridge, MA: MIT Press.

Nymoen, K., Godøy, R.I., Jensenius, A.R., & Torresen, J. (2013). Analyzing correspondence between sound objects and body motion. *ACM Transactions on Applied Perception, 10*, article 9, 1-22.

Palmer, C., Koopmans, E., Carter, C., Loehr, J., & Wanderley, M. (2009). Synchronization of motion and timing in clarinet performance. In Williamon, A., Pretty, S., and Buck, R., (Eds.), *Proceedings of the International Symposium on Performance Science*, (pp. 159–164). European Association of Conservatories (AEC), Utrecht.

Pyper, B. J. & Peterman, R. M. (1998). Comparison of methods to account for autocorrelation in correlation analyses of fish data. *Canadian Journal of Fisheries and Aquatic Sciences, 55*, 2127–2140.

Repp, B. H. & Su, Y.-H. (2012). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin & Review, 20*, 403-452.

Sherry, A. & Henson, R. K. (2005). Conducting and interpreting canonical correlation analysis in personality research: A user-friendly primer. *Journal of Personality 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.

Toiviainen, P., Luck, G., & Thompson, M. R. (2010). Embodied meter: Hierarchical 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, & Psychophysics, 76*, 591–604.

Wanderley, M. M., Vines, B. W., Middleton, N., McKay, C., & Hatch, W. (2005). The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research, 34*, 97–113.