

PERCEPTION OF EXPRESSION IN CONDUCTORS' GESTURES: A CONTINUOUS RESPONSE STUDY

GEOFF LUCK, PETRI TOIVIAINEN,
& MARC R. THOMPSON
University of Jyväskylä, Jyväskylä, Finland

THE VISUAL CHANNEL HAS BEEN SHOWN TO BE MORE informative than the auditory channel in perceptual judgments of a performer's level of expression. Previous work has revealed a positive relationship between amplitude of music-related movement and ratings of expression, for example, and observers have been shown to be sensitive to kinematic features of music-related movement. In this study, we investigate relationships between the kinematics of a conductors' expressive gestures and ratings of perceived expression. Point-light representations (totalling 10 minutes) of two professional conductors were presented to participants who provided continuous ratings of perceived valence, activity, power, and overall expression using a virtual slider interface. Relationships between these ratings and 11 kinematic variables computationally extracted from the movement data were subsequently examined using linear regression. Higher levels of expressivity were found to be conveyed by gestures characterized by increased amplitude, greater variance, and higher speed of movement.

Received November 23, 2009, accepted April 7, 2010.

Key words: expressiveness, movement kinematics, regression, continuous response, point-light display

PREVIOUS RESEARCH HAS HIGHLIGHTED relationships between corporal behavior and perception of musical expression. It has been shown, for example, that observers are able to perceive a musician's expressive intention through the visual mode alone (Davidson, 1993, 1995), while expressive body movement has been found to influence observers' level of interest in the performance (Broughton & Stevens, 2008). Expressive body movement can also be a determining factor in discriminating between specific emotional intentions (Dahl & Friberg, 2007), while a musical structure that

represents musical form, such as phrase structure, can also be perceived through movement alone (Krumhansl & Schenck, 1997). In addition, emotional characteristics of dancers can be perceived in their movements (e.g., Dittrich, Troscianko, Lea, & Morgan, 1996).

In addition to identifying specific emotions, observers can also perceive varying *magnitudes* of expression through the visual channel alone. While the ability to discriminate between different magnitudes of musical expression had already been studied in the auditory mode (Kendall & Carterette, 1990), Davidson (1993) reported that the visual channel also played a significant role in recognizing different performance manners. Further, Davidson found that for nonmusicians, it was the visual mode that most clearly specified the different expressive manners (deadpan, projected, and exaggerated).

The way the body conveys different magnitudes of expression may be related to the amplitude of the gestures produced. Davidson (2007) studied a video of a pianist performing a Beethoven Bagatelle in the deadpan, projected, and exaggerated expressive intentions. She identified expressive locations in which the same body movements in the head and hands were repeated throughout every intention. Although the movements were similar in shape throughout all performance manners, the movements were smallest in the deadpan intention and largest in the exaggerated intention.

In a related study, two of the present authors investigated whether playing with different magnitudes of expression would result in different amounts of body movement (Thompson & Luck, 2010). Pianists performed in multiple sessions an excerpt from the same piece using three different levels of expression ("minimum", "normal", and "maximum"), and their movements were recorded using an optical motion capture system. Data showed that the head and shoulders traveled a further distance overall, and showed bigger differences between expressive intentions, compared to the fingers, wrists and lower back, indicating a positive relationship between range of motion of a body part and its use in conveying expression.

Changes in amplitude of movement, then, may be a cue that observers use when judging the expressivity of a performance. An additional hypothesis is that observers react to more detailed features of the performer's movements; that is, the movement kinematics. It has been shown, for instance, that observers are sensitive to the movement kinematics of conductors' time-beating gestures. Several studies have identified the features of conducting gestures that offer cues for synchronization, employing a range of paradigms to do so, including tapping (Luck & Sloboda, 2008, 2009), temporal adjustment (Luck, 2008), and live performance (Luck & Toiviainen, 2006).

In Luck and Sloboda's (2008) study, participants synchronized finger taps with point light stimuli derived from simple conducting gestures. A series of linear regression analyses identified acceleration along the trajectory as the main cue for synchronization, with this movement feature accounting for up to 68% of the variance in participants' responses. A follow-up study, in which participants synchronized with extended sequences of beats in the form of traditional three-beat patterns (Luck & Sloboda, 2009) again identified absolute acceleration along the trajectory as the main cue for synchronization. A study in which participants adjusted the timing of an auditory stimulus until it was perceived as being in synchrony with a series of point-light conducting gestures identified acceleration along the trajectory as the sole cue for synchronization (Luck, 2008). In a more ecological study (Luck & Toiviainen, 2006), a series of cross-correlations identified acceleration along the trajectory as the primary feature with which ensemble musicians synchronized with a conductor in a live setting, while also showing that the ensemble tended to lag behind the conductor. In several of the above studies, speed of movement was also found to be a relevant kinematic feature, but less so than acceleration.

Given, then, that observers are sensitive to certain kinematics features of temporal conducting gestures, might they also be sensitive to the kinematics of more expressive gestures? The aim of the present study was to collect some preliminary data on this issue. We investigated relationships between the kinematics of conductors' expressive gestures and ratings of perceived expression in order to identify whether expression is conveyed simply by amplitude of the gestures, or if observers are, in addition, sensitive to more subtle features of the movements. Since the expression conveyed in conductors' gestures does not remain static, but rather changes and evolves over time, we employed a continuous response paradigm. This methodology is used widely in studies of emotion and music because of its dynamic nature of data collection – participants' responses can be collected throughout

the duration of a temporally extended stimulus, such as a conductor conducting a musical passage.

We combined the continuous response paradigm with a computational feature-extraction approach in which selected movement features were extracted from the gestures at regular intervals, and relationships between these features and the continuous ratings subsequently investigated. Two of the present authors have used this combination of methods in the investigation of emotional responses to musical features in music therapy improvisations (Luck et al., 2008), and in diagnosing level of mental retardation from music therapy improvisations (Luck et al., 2006). While these two studies examined perception of various types of auditory (as opposed to movement) features, the underlying principle of examining relationships between continuous ratings and computationally extracted features was the same. Wöllner and Auhagen (2008) have employed the continuous response paradigm in investigating conductors' expressive gestures, but without detailed kinematic analysis of the gestures themselves. Instead, video recordings were analyzed using the EyesWeb (Camurri & Volpe, 2004) Quantity of Motion (QoM) module, and relationships between amount of movement and ratings of expression examined.

The movement features extracted in the present study related to position, speed, acceleration, and jerk of different body parts. The first three of these were included because of their relevance in previous studies (position indicating magnitude of movement, and speed and acceleration being important in temporal conducting gestures). Jerk, meanwhile, was also included because of the potential relationship between jerkiness or smoothness of movement and emotional expression.

The present study adopted the dimensional concept of emotion in order to investigate emotional expression in a more detailed fashion (as opposed to asking observers to rate only perceived expression as a single concept). It has been suggested that the dimensional concept of emotion is particularly well-suited to studies that examine the dynamic changes in music-related emotional expression (Juslin & Sloboda, 2001). Dimensional theories of emotion hold that emotional meaning can be described within a multidimensional emotion space comprised of a small number of dimensions, most frequently cited as relating to valence, activity, and power (e.g., Osgood, Suci, & Tanenbaum, 1957). Each dimension is assumed to be anchored by semantic terms representing polar opposites, such as happy–sad for valence, active–inactive for activity, and weak–strong for potency. Schubert (2001) notes that references to the first two of these dimensions are frequently found in the music-emotion literature. Power, however, is somewhat less frequently described, and its

role in the emotion space not so clearly defined. Nonetheless, the concept of power seems apt for the description or experience of conducting gestures since physical movement can be easily described as being "powerful."

In the present study, participants were presented with point-light displays of conducting gestures, and asked to provide continuous ratings of perceived emotional expression on four scales: *valence*, *activity*, *power*, as well as overall amount of *expression*. In line with the findings of Davidson (2007) and Thompson and Luck (2010), it was predicted that ratings of perceived expression would be positively related to gesture amplitude. It was also expected that the different scales would tease out different combinations of movement features, revealing more specific relationships between movement features and ratings of expression. One might suppose, for example, that activity would be positively related to at least movement amplitude and speed, and that power would be positively related to movement amplitude and perhaps jerk. Precise relationships between the different scales and movement features were, however, hard to predict given the exploratory nature of this study. Nonetheless, it was further predicted that ratings of expression would be more strongly associated with movement features of the left hand given the traditional role of the left and right hands in conducting. In other words, while the right hand is primarily reserved for time-keeping duties, it is the left hand that is traditionally used to convey expression (Rudolf, 1980). Thus, we would expect the left hand to be more important to observers when rating the expressiveness of conducting gestures.

Method

Participants

Twenty-four individuals (mean age = 24.6 years, $SD = 2.7$, females = 16) provided continuous ratings of perceived expression. All participants were students attending the University of Jyväskylä, Finland. While a majority of the participants had received five or more years of music training (54%), only four participants had any formal conducting training. The participants were recruited using a university mailing list for students and were compensated for their time with a free cinema pass.

Gesture Recording and Stimulus Generation

The gestures of two early-career professional conductors (both enrolled on the conducting course at the Sibelius Academy, Finland) were recorded with an optical motion-capture system (Qualisys ProReflex) while they

directed an ensemble of approximately 40 instrumentalists and singers. The piece being rehearsed was Mozart's Requiem Mass in D minor (K. 626). The three dimensional position of eight reflective markers attached to the hands (base of middle finger), wrists, elbows, and shoulders of each conductor was tracked at 60 Hz. Six excerpts (totaling 10 minutes) covering a range of expression and movement types (i.e., varying in terms of movement amplitude, speed, smoothness, etc.) were selected as stimuli, and the corresponding movement data then transformed into six QuickTime movies. To create the QuickTime movies, the movement data were imported into MATLAB, where animation frames of the point-light displays were created using the animation-related functions included in the MoCap Toolbox (Toivainen, 2008). Each animation cell was saved as a JPEG file and then opened in QuickTime 7 as an image sequence and saved as a self-contained QuickTime movie. The animations had a frame rate of 30 Hz and segments were created to connect relevant markers to each other.

Procedure

A Max/MSP patch was created to play back the QuickTime movies, provide the slider interface that participants used to give their ratings, and save the ratings to file. A screen shot of the patch is shown in Figure 1. The six stimuli were presented in four identical blocks, one each for ratings of perceived activity, valence, strength, and overall expression. Within-block presentation order was randomized each time. Further, the order in which the blocks were presented was unique for each of the 24 participants. This was possible because four blocks could be ordered in a maximum of 24 unique permutations. Each participant initially practiced using the interface by providing continuous ratings for stimuli similar to the experimental stimuli. Once the participant expressed confidence that he/she understood the task, the experimenter left the room. Each block was prefaced by instructions on how each perceptual dimension should be rated and that the participant should use the full range of the slider when making perceptual judgements. The slider was controlled by clicking and dragging with the mouse. The instructions for each block were as follows:

Expression. Move the slider to the right when you feel the gestures are more expressive and move it to the left when you feel they are less expressive.

Valence. Move the slider to the right when you feel the gestures are more pleasant and move it to the left when you feel they are less pleasant.

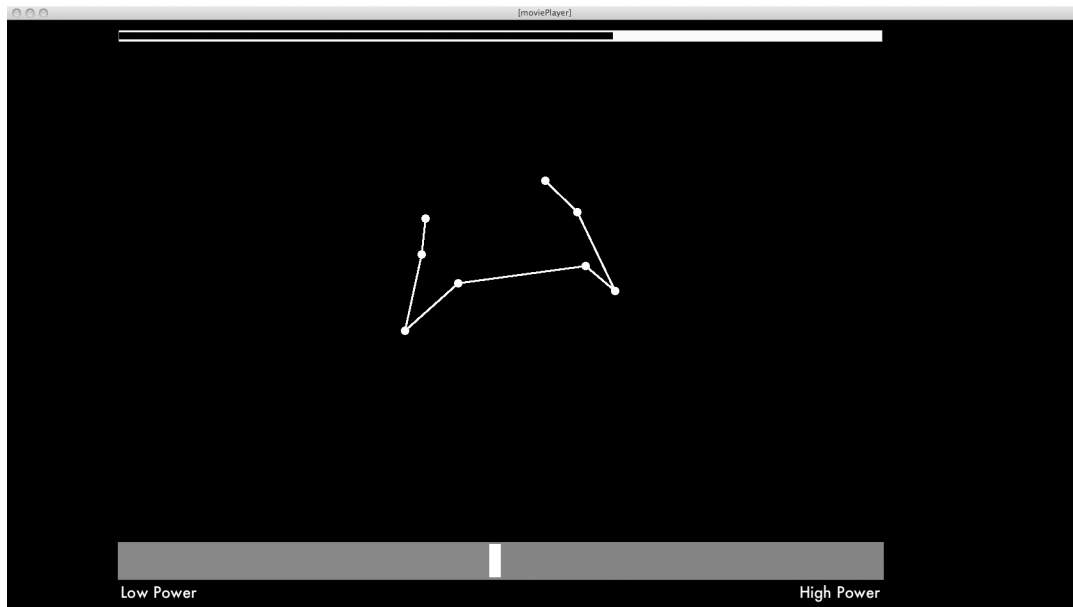


FIGURE 1. A screenshot of the Max/MSP patch displaying the stimuli. The user watches the point-light display while dragging the vertical white bar across the horizontal grey bar as their perception of expression, power, valance, and activity changes. A yellow status bar at the top of the screen indicates the temporal progression of the animation.

Activity. Move the slider to the right when you feel the gestures have a high level of activity and move it to the left when you feel they have a low level of activity.

Power. Move the slider to the right when you feel the gestures are more powerful and move it to the left when you feel they are less powerful.

Within a block, the QuickTime movies were played back automatically with an interval of 6 s between each. The positions on the slider ranged from -100 to $+100$. The Max/MSP patch recorded the position of the slider at a frequency of 4 Hz, allowing a high level of accuracy and detection of abrupt changes in the slider's position. Once the participant had been exposed to an entire block, they were given the option to take a break before starting the following block. Data collection lasted approximately 45 minutes per participant.

Movement Feature Extraction

In order to examine relationships between the perceptual ratings and the conductors' movements, 11 hand-related movement features were extracted from the raw movement data.

The first four movement features were simply the position trajectories relating to the horizontal (x-axis) and vertical (y-axis) movement of the conductors' hands.

Movement depth (z-axis) was not considered for analysis as the participants were only exposed to animations displaying horizontal and vertical movements.

The fifth feature was a single vector representing the frame-by-frame distance between the hand markers. This feature was selected for analysis as one might assume that the distance between the hands was a distinguishing feature of large and small movements and would have influenced participants' perceptual ratings.

Features six to eleven were related to the kinematics of the hand movements. The absolute values of velocity, acceleration, and jerk were estimated along the horizontal and vertical dimensions of the movement data by applying numerical differentiation to the position data. The time derivative at each frame was estimated using a Savitzky-Golay FIR smoothing filter, which fitted a second order polynomial to seven consecutive frames centred at the frame in question and computed the derivative of the thus obtained polynomial at the centre frame (Luck & Toiviainen, 2006). From these column vectors, the instantaneous velocity, acceleration, and jerk were derived as the length of the vector consisting of the components of velocity, acceleration, and jerk respectively.

The final step was to make all of the movement variable column vectors the same length as the slider position vectors. This was achieved by taking the average of 30 consecutive frames occurring prior to each reading of the slider, thus down-sampling the movement data

from 60 Hz to 4 Hz. This averaging method allowed for direct examination between the slider position ratings and the average of the movement data occurring 250 ms prior to the rating.

To sum up, the eleven movement features from each conducting performance were the horizontal and vertical positional data of each hand (4), the distance vector between each hand (1), and the instantaneous velocity, acceleration and jerk of each hand marker (6).

Results

Calculating Mean Ratings

Mean ratings (in the form of mean time series) were first calculated for each excerpt. Each participant's individual rating was then correlated with each of these means, and participants with either nonsignificant or negative coefficients were regarded as outliers and removed from further analyses. Mean ratings were then recalculated from remaining participants' ratings. For the six different stimuli, an average of 4.6 participants were removed from the initial expression ratings, 7.2 from the valence ratings, 3 participants from the activity ratings, and 5 participants from the power ratings.

Next, in order to see how ratings of the four dimensions related to each other, interdimension correlations, based upon the mean rating for each of the four dimensions, were calculated. The correlation matrix for the averaged data is shown in Table 1 and reveals that ratings for the four dimensions were generally moderately correlated with each other, except for activity and power which were quite highly correlated. The lowest correlation was between expression and valence, and the highest correlation was between activity and power.

TABLE 1. Correlation Matrix Showing the Interdimension Correlations Between Expression, Valence, Activity, and Power.

	Expression	Valence	Activity	Power
Expression	1.00	.37	.45	.49
Valence	—	1.00	.55	.43
Activity	—	—	1.00	.79
Power	—	—	—	1.00

Note: All coefficients were significant at $p < .01$, and N for all time-series was 2222.

Lag Analysis

Mean ratings for each excerpt were then crosscorrelated with the corresponding movement variables. In each case, the maximal cross-correlation within the range -20 samples to $+20$ samples (-5 s to $+5$ s) indicated the lag between the musical feature and participants' response to it. Variables were subsequently time shifted so that the peak correlation occurred at lag = zero. This was done separately for each of the 24 mean ratings (six excerpts \times four dimensions), and it was these final time-shifted datasets that were used in subsequent regression analyses. Table 2 shows the mean lag for each of the 11 movement features for each rating scale.

Regression Analysis

To model the experimental data, we employed ordinary least squares linear regression, in which the movement features were entered as predictors of participants' emotion ratings. Excerpts 1–3 were produced by Conductor 1, while excerpts 4–6 were produced by Conductor 2. Preliminary inspection of the data revealed some differences in terms of relationships between ratings of perceived emotion and their movement features such that

TABLE 2. Mean Lags (s) for Each of the 11 Movement Features and the Four Rating Scales. Positive Values Indicate the Amount of Time Elapsed Between a Movement Feature Occurring and Participants Responding to It.

Movement feature	Expression	Valence	Activity	Power	Overall
Distance between hands	3.21	1.96	2.33	2.54	2.51
Right hand x position	0.58	0.79	1.50	2.08	1.24
Right hand y position	1.38	1.38	1.33	1.46	1.39
Right hand speed	1.88	2.04	1.33	3.25	2.13
Right hand acceleration	2.42	2.25	3.04	2.33	2.51
Right hand jerk	2.29	1.50	1.79	1.67	1.82
Left hand x position	2.79	1.71	1.58	1.46	1.89
Left hand y position	2.29	1.54	1.42	1.42	1.67
Left hand speed	2.79	1.63	1.71	1.33	1.87
Left hand acceleration	2.71	1.33	1.54	1.21	1.70
Left hand jerk	2.67	1.58	1.83	1.92	2.00

TABLE 3. Summary of the Eight Regression Models.

Dimension	Conductor	F ratio	df	R ²	R ² _{adj}
EXPRESSION	Both	260.71	11,2231	.562	.560
	1	105.57	11,1161	.500	.495
	2	160.61	11,1058	.625	.622
VALENCE	Both	314.18	11,2244	.606	.604
	1	173.79	11,1169	.621	.617
	2	226.53	11,1063	.701	.698
ACTIVITY	Both	282.38	11,2253	.580	.578
	1	174.62	11,1173	.621	.617
	2	284.24	11,1068	.745	.743
POWER	Both	265.42	11,2222	.568	.566
	1	118.89	11,1154	.531	.527
	2	272.42	11,1056	.739	.737

Note: All F ratios were significant at $p < .001$.

analyzing all six excerpts as a whole could obscure true effects. Thus, in addition to analyzing all excerpts together, excerpts 1–3 (Conductor 1) and excerpts 4–6 (Conductor 2) were pooled and analyzed separately. In order to examine larger structures and overall themes in terms of relationships between movement features and ratings of expression, all variables were smoothed using a running mean with a 10-point-span (equivalent to 2.5 s) prior to analysis.

Twelve separate linear regression analyses were carried out, one for each of the four emotion dimensions on: (a) the whole dataset, (b) pooled data for excerpts 1–3, and (c) pooled data for excerpts 4–6. In each

TABLE 4. Regression Models for EXPRESSION.

Movement feature	Beta value		
	Both	Conductor 1	Conductor 2
Distance between hands	.006 [†]	.082***	-.094**
Right hand x position	.024 [†]	.094***	.013 [†]
Right hand y position	.109***	.027 [†]	.065**
Right hand speed	.002 [†]	.607***	-.057 [†]
Right hand acceleration	.664***	.326**	.883***
Right hand jerk	.052	-.293***	.032 [†]
Left hand x position	-.142***	-.172***	-.085***
Left hand y position	.024 [†]	.145***	.143***
Left hand speed	-.001 [†]	-.075 [†]	-.024 [†]
Left hand acceleration	.696***	.176**	.900***
Left hand jerk	-.695***	-.057 [†]	-.894***

Note: Beta values indicate the strength and direction of the relationship between each movement feature and the mean expression rating; * $< .05$, ** $< .01$, *** $< .001$, [†]not significant.

TABLE 5. Regression Models for VALENCE.

Movement feature	Beta values		
	Both	Conductor 1	Conductor 2
Distance between hands	-.007 [†]	.023 [†]	.050 [†]
Right hand x position	.040 [†]	.085***	-.029 [†]
Right hand y position	.134***	.165***	.063**
Right hand speed	.803***	.577***	.584***
Right hand acceleration	-1.407***	-.115 [†]	-1.138***
Right hand jerk	.987***	.149**	.855***
Left hand x position	-.026 [†]	-.071**	.091***
Left hand y position	.008 [†]	.259***	-.029 [†]
Left hand speed	1.040***	.422***	1.108***
Left hand acceleration	-1.745***	-.253**	-1.683***
Left hand jerk	1.103***	.127**	1.028***

Note: Beta values indicate the strength and direction of the relationship between each movement feature and the mean valence rating; * $< .05$, ** $< .01$, *** $< .001$, [†]not significant.

analysis, the 11 movement variables were entered simultaneously. Significant models emerged for all twelve analyses, and are summarized in Table 3. Beta values and significance levels for the movement variables are shown in Tables 4–7.

Expression. It can be seen from Table 4 that, when both conductors were analyzed together, ratings of perceived expression were positively related to right-hand y position (greater perceived expression when the right hand was held higher), and acceleration of both hands, and negatively related to left-hand x position (greater perceived

TABLE 6. Regression Models for ACTIVITY.

Movement feature	Beta values		
	Both	Conductor 1	Conductor 2
Distance between hands	.082***	.104***	.129***
Right hand x position	.020 [†]	.056**	-.007 [†]
Right hand y position	.108***	.107***	.015 [†]
Right hand speed	.249***	.735***	-.312***
Right hand acceleration	.032 [†]	-.031 [†]	1.247***
Right hand jerk	.291***	.031 [†]	-.353***
Left hand x position	-.083***	-.032 [†]	-.034 [†]
Left hand y position	-.011 [†]	.276***	.013 [†]
Left hand speed	.378***	.049 [†]	.319***
Left hand acceleration	-.307***	.002 [†]	-.124 [†]
Left hand jerk	.127*	.092*	-.015 [†]

Note: Beta values indicate the strength and direction of the relationship between each movement feature and the mean activity rating; * $< .05$, ** $< .01$, *** $< .001$, [†]not significant.

TABLE 7. Regression Models for POWER.

Movement feature	Beta values		
	Both	Conductor 1	Conductor 2
Distance between hands	.036 [†]	-.011 [†]	.149***
Right hand x position	-.019 [†]	.058**	-.072***
Right hand y position	.168***	.185***	.107***
Right hand speed	.927***	.617***	.995***
Right hand acceleration	-1.554***	-.126*	-.1487***
Right hand jerk	.937***	.041 [†]	.922***
Left hand x position	-.061***	-.027 [†]	-.038 [†]
Left hand y position	-.040*	.119***	.028 [†]
Left hand speed	.736***	.162***	.652***
Left hand acceleration	-.949***	-.046 [†]	-.939
Left hand jerk	.614***	.174***	.665***

Note: Beta values indicate the strength and direction of the relationship between each movement feature and the mean power rating; * < .05, ** < .01, *** < .001, [†]not significant.

expression when the left hand was held closer to the body) and jerk.¹ When Conductor 1 was analyzed separately, ratings of perceived expression were positively related to distance between the hands, right-hand x position, speed, and acceleration, and left-hand y position and acceleration, and negatively related to right-hand jerk, and left-hand x position. Regarding Conductor 2, ratings of perceived expression were positively related to right-hand y position and acceleration, and left-hand y position and acceleration, and negatively related to distance between the hands, and left-hand x position and jerk.

Valence. Table 5 reveals that, when both conductors were analyzed together, ratings of perceived valence were positively related to right-hand y position, speed, and jerk, as well as left-hand speed and jerk, and negatively related to both right- and left-hand acceleration. For Conductor 1, meanwhile, perceived valence was positively related to right-hand x position, y position, speed, and jerk, and left-hand y position, speed, and jerk, and negatively related to left-hand x position and acceleration. For Conductor 2, perceived valence was positively related to right-hand y position, speed, and jerk, and

left-hand x position, speed, and jerk, and negatively related to both right and left-hand acceleration.

Activity. Table 6 shows that, when both conductors were analyzed together, ratings of perceived valence were positively related to distance between the hands, right-hand y position, speed, and jerk, and left-hand speed and jerk, and negatively related to left-hand x position and acceleration. For Conductor 1, ratings of perceived activity were positively related to distance between the hands, right-hand x position, y position, and speed, and left-hand y position and jerk. Regarding Conductor 2, perceived activity was positively related to distance between the hands, right-hand acceleration, and left-hand speed, and negatively related to right-hand speed and jerk.

Power. Table 7 shows that, when both conductors were analyzed together, ratings of perceived power were positively related to right-hand y position, as well as speed and jerk of both hands, and negatively related to left-hand x position, and acceleration of both hands. For Conductor 1, meanwhile, perceived power was positively related to right-hand x position, y position, and speed, and left-hand y position, speed, and jerk, and negatively related to right-hand acceleration. For Conductor 2, perceived power was positively related to distance between the hands, right-hand y position, speed, and jerk, and left-hand speed and jerk, and negatively related to right-hand x position and acceleration.

To help visualize the success of the 12 models in predicting participants' ratings, the predicted values of expression, valence, activity, and power resulting from these models were plotted against the actual mean rating for each of these dimensions, for each of the six excerpts. These plots are shown in Figure 2. It can be seen that the predicted ratings generally correspond quite closely to the actual mean ratings. The fit between predicted and actual ratings, however, varies from excerpt to excerpt, and, since the different models explain different amounts of variance, between both the conductors and the four dimensions.

Finally, potential issues of multicollinearity were investigated in order to check the accuracy and stability of the models. An examination of two indices of multicollinearity revealed no serious concerns related to this phenomena. More specifically, mean variance inflation factors (VIFs)—which indicate whether a predictor has a strong linear relationship with the other predictors—for the expression, valence, activity, and power models (by conductor) were 6.5/11.2, 7.2/27.1, 6.2/13.3, and 3.4/34.9 respectively. These figures suggest that small levels of multicollinearity may be present (see Bowerman & O'Connell, 1990). However, tolerances for most variables were at least .1, and in the majority of cases were between .4 and .8. Tolerance is the percentage of

¹The coordinate system was such that larger values on the y axis indicated higher vertical position, while larger values on the x axis indicated a greater degree of movement towards the left of the conductors' body. Consequently, a positive relationship between each dimension's rating and y and x position indicated that, when the rating had a higher value, position was high and shifted towards the left (e.g., high up, and further away from the body for the left hand but closer to the body for the right hand), and vice versa for a negative relationship (e.g., low down, and closer to the body for the left hand but further away from the body for the right hand).

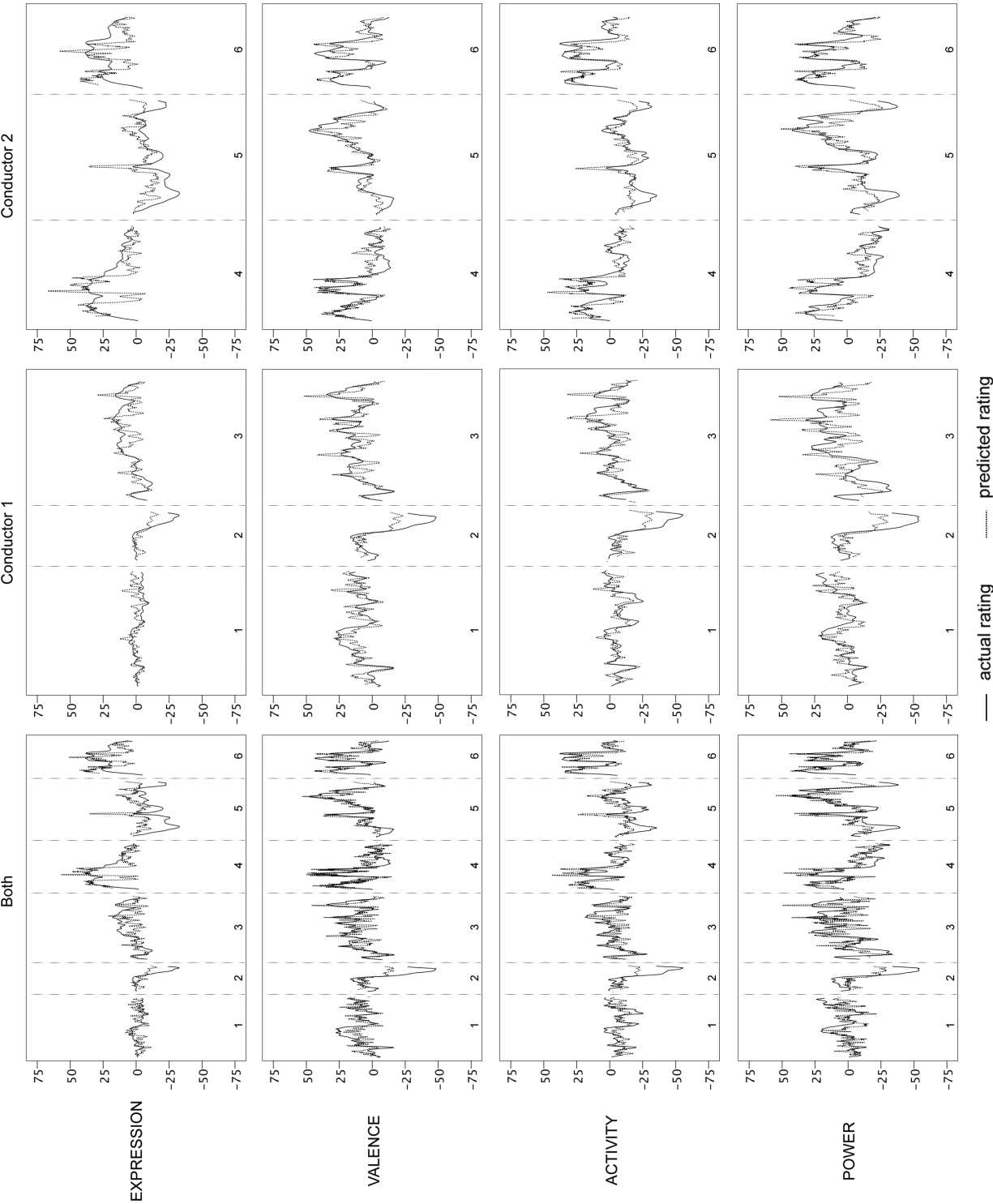


FIGURE 2. Actual and predicted ratings (vertical axes) of each of the four emotion dimensions plotted again time (horizontal axes) for all six excerpts.

the variance in a given predictor that cannot be explained by the other predictors (values should be multiplied by 100 to obtain the percentage). Only values below .1 indicate serious problems (see Field, 2005), and the fact that most of the variables had high tolerances indicated no problems of multicollinearity. In sum, then, we were confident that multicollinearity was not unduly affecting the accuracy or stability of the regression models.

Discussion

Overall, the regression models accounted for between 50 and 75 percent of the variance in participants' ratings of the four dimensions (mean variance = 56% for expression, 64% for valence, 65% for activity, and 61% for power). When both conductors were analyzed together, gestures rated as being more expressive overall were characterized by vertically extended and accelerated right-hand movement, and smooth but accelerated, and less horizontally extended, left-hand movement. Gestures receiving higher ratings of valence were characterized by vertically extended right-hand movement, and less accelerated but fast and jerky movement of both hands. Gestures receiving higher ratings of activity were characterized by vertically extended right-hand movement, greater distance between the hands, fast and jerky movement of both hands, and less horizontally extended and accelerated left-hand movement. Finally, gestures rated as being more powerful were characterized by vertically extended right-hand movement, less accelerated but fast and jerky movement of both hands, and both horizontally and vertically less extended left-hand position. These descriptions are summarized in Table 8.

A few points regarding these descriptions are worth noting. First, the similarity between the movement descriptions for activity and power, and the dissimilarity between expression and valence, mirror the highest and lowest inter-dimension mean rating correlations shown in Figure 1 above. Second, while the right hand is vertically extended for all dimensions (likely due to it being

consistently held high to convey primarily temporal information using traditional time-beating gestures), the beta value for this feature was highest for power, indicating that this feature was most strongly associated with conveying power. Third, speed, acceleration, and jerk appear to be important for all four measured dimensions, the direction of the relationships being identical for all three components of Osgood et al.'s (1957) dimensional theory (valence, activity, and power). Fourth, and related to the previous point, hand position seems to be the feature that differentiates between Osgood et al.'s dimensions, activity being represented in the same way as valence but with hands held further apart, and power further specifying the right hand be held highest, and the left hand close to the body. Finally, these results provide partial support for the hypothesis that higher ratings of expression will be related to greater movement amplitude inasmuch as higher ratings of activity were related to the hands being held further apart, and higher ratings of power were related to the right hand being held the highest.

One problem with generalizing the results this way is that it can obscure stylistic differences. That is, if we look for commonalities between the conductors, we may miss more important relationships between the movement features and ratings of expression that are conductor-specific. An alternative approach, therefore, is to look at each conductor individually.

Regarding perceived expression, the main differences in terms of significant movement features between the conductors were the direction of the relationship with distance between the hands (positive for Conductor 1, negative for Conductor 2), horizontal (Conductor 1) versus vertical (Conductor 2) extension of the right hand, higher speed and smoother movement of the right hand for Conductor 1, and smoother left-hand movement for Conductor 2. For perceived valence, the main differences were horizontally extended right-hand for Conductor 1, significantly less accelerated right-hand movement for Conductor 2, the direction of the relationship with horizontal extension of the left hand (negative for Conductor 1, positive for Conductor 2), and the significance of vertically extended left-hand position. Regarding perceived activity, the main differences were more horizontally and vertically extended right hand position for Conductor 1, the direction of the relationship with right-hand speed (positive for Conductor 1, negative for Conductor 2), more accelerated but smoother right-hand movement for Conductor 2, vertically extended left-hand position for Conductor 1, and left-hand speed for Conductor 2. Finally, for perceived power, the main differences were distance between the hands for Conductor 2, the direction of the relationship

TABLE 8. The Primary Movement Features Associated with Mean Ratings of Expression, Valence, Activity, and Power When Both Conductors Were Analyzed Together.

Dimension	Movement features
Expression	Accelerated but smooth
Valence	Fast, jerky, but less accelerated
Activity	Hands further apart, fast and jerky, less accelerated
Power	Fast and jerky, less accelerated, right hand held highest, left hand close to body

with horizontal extension of the right-hand (positive for Conductor 1, negative for Conductor 2), smoother movement for Conductor 2, and vertically extended left-hand position for Conductor 1.

It is interesting that most of these differences relate to hand position, with fewer differences between the conductors in terms of speed, acceleration, and jerk of movement. It seems, therefore, that differences in hand position help differentiate not only between different expressive dimensions, but also between different conductors. It also seems that the relatively stable relationships between speed, acceleration, and jerk, and the expression dimensions seen when both conductors were analyzed together break down when the conductors are analyzed separately.

In addition, it is worth noting that, in several cases, the direction of the relationship between movement features and the expressive dimensions are at odds with each other (such as distance between the hands being positively related to the expression dimension for Conductor 1, but negatively related for Conductor 2). These differences have the effect of canceling each other out when both conductors are analyzed together, and thus limits the extent to which the results of the overall analysis (both conductors together) can be generalized.

Analyzing the results by conductor, then, helps to reveal stylistic differences in conducting technique. Moreover, Table 3 reveals that the regression models for Conductor 2 consistently “outperformed” those for Conductor 1: mean variance explained for Conductor 1 was 57%, as compared to 71% for Conductor 2. This implies that relationships between movement features and perceived expression were clearer for Conductor 2 than Conductor 1, and suggests that Conductor 2 was able to express emotion in a more consistent manner.

Both stylistic differences between conductors, as well as more general relationships between movement features and perceived expression, might be clarified in future work by replicating this study with more conductors. The investigation of two conductors in the present study reduces the extent to which our results can be generalized, particularly in light of the idiosyncrasies they demonstrated.

Another interesting approach would be to have the conductors rate each other (and perhaps themselves) in terms of perceived expression. This might be further extended by having other skilled conductors rate these conductors. Since musicians with previous conducting experience have been shown to synchronize more consistently with temporal conducting gestures compared to musicians with synchronization experience only or non-musicians (Luck & Nte, 2008), one might speculate that a similar effect of domain-specific experience would be evident in ratings of perceived expression. That is, skilled

conductors might have a deeper understanding of how expression is conveyed by movement, and thus be able to make better use of the kinematic information available.

This also raises the question of potential connections between expressivity and timing. Might there be a connection between expressivity and timing, for example, such that observers can synchronize better with more expressive conductors? Or vice versa? There is also the issue of how the expressiveness of conductors’ gestures affects timing within an ensemble, not just between the ensemble and the conductor. One might suppose that gestures containing smaller peaks in acceleration (in the present study, those characterized by, for example, high levels of valence) would be harder to synchronize with accurately compared to gestures containing more marked peaks in acceleration (in the present study, gestures characterized by higher levels of overall expression).

Furthermore, there is the issue of the use of other body parts in expressive conducting. A number of studies have shown, for example, that the conductor’s face conveys important expression-related information (e.g., Dan, 2005; Mayne, 1992; Wöllner, 2008; Wöllner & Auhagen, 2008). Wöllner and Auhagen argue that, while a reduction of conducting gestures to the movement of a limited number of points in space (such as the tip of the baton, or the hands, for instance) allows a controlled experimental setting, validity could be increased if other aspects, such as facial expression, are included in stimuli. We agree with this assertion, but chose in the present study to focus on the more general role of gross body movements in conveying expression.

In sum, this study provides preliminary data on relationships between characteristics of conductors’ gestures and perceived expression. We found that observers (including nonmusicians) were sensitive not only to movement amplitude but also more fine-grained kinematic features when rating perceived expression. In conclusion, we found that higher levels of expressivity can be conveyed by gestures characterized by increased amplitude (greater distance between the hands, lower and more horizontally extended left-hand position, higher right-hand position), greater variance (jerkier movement of both hands), and higher speed of movement (faster movement of both hands).

Author Note

Correspondence concerning this article should be addressed to Geoff Luck, Academy of Finland Research Fellow Finnish Centre of Excellence in Interdisciplinary Music Research, Department of Music, P.O. Box 35(M) 40014, University of Jyväskylä, Finland. E-MAIL: geoff.luck@jyu.fi

References

- BOUGHTON, M., & STEVENS, C. (2008). Music, movement and marimba: An investigation of the role of movement and gesture in communicating musical expression to an audience. *Psychology of Music*, 37, 137–153.
- BOWERMAN, B. L., & O'CONNELL, R. T. (1990). *Linear statistical models: An applied approach*. Boston, MA: PWS Publishing Co.
- CAMURRI, A., & VOLPE, G. (Eds.). (2004). *Gesture-based communication in human-computer interaction*. Berlin: Springer.
- DAHL, S., & FRIBERG, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception*, 24, 433–454.
- DAN, K. (2005). Facial expression and eye contact used by instrumental conductors: Practical applications and exercises. *Dissertation Abstracts International*, 66(02), 399 (UMI No. 3166080).
- DAVIDSON, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, 21, 103–113.
- DAVIDSON, J. W. (1995). What does the visual information contained in music performances offer the observer? Some preliminary thoughts. In R. Steinberg (Ed.), *Music and the mind machine: The psychopathology of the sense of music XXX* (pp. 105–113). Berlin: Springer-Verlag.
- DAVIDSON, J. W. (2007). Qualitative insights into the use of expressive body movement in solo piano performance: A case study approach. *Psychology of Music*, 35, 381–401.
- DITTRICH, W. H., TROSCIANKO, T., LEA, S. E. G., & MORGAN, D. (1996). Perception of emotion from dynamic light-point displays represented in dance. *Perception*, 25, 727–738.
- FIELD, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.
- JUSLIN, P. N., & SLOBODA, J. A. (2001). *Music and emotion: Theory and research*. London: Oxford University Press.
- KENDALL, R. A., & CARTERETTE, E. C. (1990). The communication of musical expression. *Music Perception*, 8, 129–164.
- KRUMHANS, C. L., & SCHENK, D. L. (1997). Can dance reflect the structural and expressive qualities of music? *Musicae Scientiae*, 1, 63–83.
- LUCK, G. (2008, August). Spatio-temporal cues for synchronization with conductors' gestures. *Proceedings of the 10th International Conference on Music Perception and Cognition (ICMPC10)*, Sapporo, Japan.
- LUCK, G., & NTE, S. (2008). A new approach to the investigation of conductors' gestures and conductor-musician synchronization, and a first experiment. *Psychology of Music*, 36, 81–99.
- LUCK, G., RIIKKILÄ, K., LARTILLOT, O., ERKKILÄ, J., TOIVIAINEN, P., MÄKELÄ, A., ET AL. (2006). Exploring relationships between level of mental retardation and features of music therapy improvisations: A computational approach. *Nordic Journal of Music Therapy*, 15, 30–48.
- LUCK, G., & SLOBODA, J. A. (2008). Exploring the spatio-temporal properties of the beat in simple conducting gestures using a synchronization task. *Music Perception*, 25, 225–239.
- LUCK, G., & SLOBODA, J. A. (2009). Spatio-temporal cues for visually-mediated synchronization. *Music Perception*, 26, 465–473.
- LUCK, G., & TOIVIAINEN, P. (2006). Ensemble musicians' synchronization with conductors' gestures: an automated feature-extraction analysis. *Music Perception*, 24, 189–200.
- LUCK, G., TOIVIAINEN, P., ERKKILÄ, J., LARTILLOT, O., RIIKKILÄ, K., MÄKELÄ, A., ET AL. (2008). Modelling the relationships between emotional responses to, and musical content of, music therapy improvisations. *Psychology of Music*, 36, 25–46.
- MAYNE, R. G. (1992). An investigation of the use of facial expression in conjunction with musical conducting gestures and their interpretation by instrumental performers. *Dissertation Abstracts International*, 53(08), 2729 (UMI No. 9238229)
- OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. (1957) *The measurement of meaning*. Champaign, IL: University of Illinois Press.
- RUDOLF, M. (1980). *The grammar of conducting*. New York: Schirmer Books.
- SCHUBERT, E. (2001). Continuous measurement of self-report emotional response to music. In P. N. Juslin & J. A. Sloboda (Eds.) *Music and emotion: Theory and research* (pp. 393–414). Oxford, UK: Oxford University Press.
- THOMPSON, M. R., & LUCK, G. (2010). *Exploring relationships between pianists' body movement, their expressive intentions, and structural elements of the music*. Manuscript submitted for publication.
- TOIVIAINEN, P. (2008). *MoCap toolbox manual*. University of Jyväskylä: Jyväskylä, Finland. Available at <http://www.jyu.fi/music/coe/materials/mocaptoolbox/MCTmanual>.
- WÖLLNER, C. (2008). Which part of the conductor's body conveys most expressive information? A spatial occlusion approach. *Musicae Scientiae*, 12, 249–272.
- WÖLLNER, C., & AUHAGEN, W. (2008). Perceiving conductors' expressive gestures from different visual perspectives. An exploratory continuous response study. *Music Perception*, 26, 129–143.

