The conductor's intensity gestures

Isabella Poggi, Francesca D'Errico and Alessandro Ansani

Abstract

This work is aimed at outlining a repertoire of conductors' gestures. In this perspective, it presents two studies that investigate a specific subset of the body signals of orchestra and choir conductors, namely, the gestures for musical intensity. First, an observational qualitative study, based on a systematic coding of a corpus of fragments from orchestra concerts and rehearsals, singled out 21 gestures, in which either the gesture as a whole or some aspects of it conveyed indications for *forte*, *piano*, *crescendo*, or *diMinuendo*; some are symbolic gestures, used either with the same meaning as in everyday interaction or with one specific of conductors; others are iconic gestures, both directly or indirectly iconic. Second, in a perception study, a questionnaire submitted to 77 participants tested if 8 gestures of intensity out of the 21 singled out by the coding study are in fact shared and understood, and whether they are better interpreted by music experts than by laypeople. Results showed that the tested gestures are fairly comprehensible, not only by experts but also by non-expert participants, probably due, for some gestures, to their high level of iconicity, and for others to their closeness to everyday gestures.

Keywords

conducting, gesture, Music performance, intensity, dynaMic indication, iconic gestures, syMbolic gestures

Performance in rehearsal and in concert is produced by the musicians' body movements; yet it is not only hands pressing on keys, touching strings, or moving bows: feet, trunk, and face accompany the rhythm and express concentration or emotions. The performers' body movements, postures, even physical appearance, sometimes only display cognitive and emotional states like concentration or passion or accompany technical movements "helping" musicians to perform them more easily or effectively; and in both cases, they do not necessarily communicate anything to others. But the conductor's body movements are by definition communicative: although

not directly making music, the maestro leads and synchronizes the performers' actions, by imprinting an interpretive view on the music to be played, and communicating it to performers by means of his or her whole body.

This work is framed within a general project aimed at studying the communicative signals used by conductors, but here in particular we focus on the gestures of intensity: the hand movements providing dynamic indications such as *forte*, *piano*, *crescendo*, *diMinuendo*, *pianissiMo*, and *sforzato*.

Why study the movements of conducting, to the point of writing down a whole repertoire of them? The first objective of this endeavor is a genuinely theoretical one: this work is framed within a general body of research aimed at demonstrating that for body communication systems, as for verbal languages, it is possible to write down lexicons, that is, lists of systematic correspondences between signals and meanings (Poggi, 2007; Poggi & D'Errico, 2010). Previous research has shown the possibility of writing down a gestionary (lexicon of gestures) of Italian symbolic gestures, a gazeionary (lexicon of gaze), and a touchionary (lexicon of touch), thus describing in a lexicographic fashion some body communicative systems used in everyday life. In this perspective, our aim, when it comes to the "language of conducting," is to investigate if the communication systems used in music performance are such stuff as everyday communication is made of, or if they make up a specific jargon, and if it is possible to find out similar or different devices in them. On the applicative side, studying the forms and meanings of conductors' gestures and their processes of production and comprehension may have a twofold pedagogical import: deepening the conductors' awareness of their communicative instruments may enhance the teaching of conducting, and providing new tools for conducting effectiveness may have an impact on choir and ensemble practice. Finally, on the technological side, to have a more systematic knowledge of conducting signals may also bear on the building of virtual conducting devices, among which are conductor embodied agents and robots (see ABB's YuMi conductor robot).

In the following, after overviewing previous work on the role of the body in music performance and introducing the idea of the conductor's communicative system, we present a coding study that singled out the conductor's gesture of intensity, and a perception study on their recognition by participants expert versus non-expert in music.

The body in music performance

Ever since the seminal works by Davidson and other authors (Davidson, 1991, 1993; Davidson & Correia, 2002), scholars in the psychology of music have realized the importance of analyzing body movement to give a complete picture of the cognitive and affective processes implied in music performance and perception. The literature in this field has focused on five main topics: first, the ways in which performers' movements influence the perception and the subjective experience of music. For example, Davidson (1993, 1994, 2012) presented the same piece played with three different expressive manners with vision alone, sound alone, or with both together, and found out that vision provides information about expressive intention even better than the other modes, further demonstrating that information about structural features of a music piece is also communicated to observers through body movement; and investigating the relationship between facial and body movement in a clarinet duo and in a pianist she highlighted that in both, a core swaying movement is a fundamental element of movement expression and that the combination of body and facial expression provides information about the meaning and articulation of musical structures. According to Vines, Krumhansl, Wanderley, and Levitin (2006) who study the dynamics of sensory integration in the perception of

musical performance, auditory and visual channels convey similar experiences of phrasing but different experiences of tension as a cue to expressivity, and the musicians' movements, extending the sense of phrasing, indicate musical interpretation and anticipate changes in emotional content. Demos, Chaffin, and Kant (2014) use detrended fluctuation analysis to identify synergies of body parts, for example, when the swaying of the body congruently accompanies rhythm, melody, and musical structure, and showed they are affected by the performer's expressive intentions.

A second research area concerns the functions of performers' movements. Wanderley, Vines, Middleton, McKay, and Hatch (2005) and Wanderley and Vines (2006) showed that clarinetists' ancillary gestures, swaying movements not directly involved in the production of sound, inform not only about the melody, rhythm, and articulation of music, but also about the emotions felt by performers and intended by composers (on the same issue, interesting treatments can be found in Juslin, 2005; Juslin & Timmers, 2010); they are consistent across standard, expressive, and immobilized manners of performance, and strongly related to the score, the different styles of expressive movements across performers, and the audience's perception.

Poggi (2006) showed that during piano performance, the pianists' head, mouth, and gaze movements can express cognitive processes of memory or attention, emotions, musical structure, but also simply collaborate with the motor processes required by the technical gesture; so it is possible to outline a lexicon of the pianist's gaze, head, and body. After analysis of different performances of a Beethoven bagatelle, Davidson (2007) discovered specific expressive locations and movement shapes in the pianist's body, the former common but the latter used flexibly across manner and time. Using motion tracking, Thompson and Luck (2011) found that shoulders move more than fingers, wrists, and lower back, and their quantity of movement change from expressive to non-expressive performance.

A third set of studies concerns body communication among co-performers. According to Williamon and Davidson (2002), through negotiation and coordination during rehearsals, specific gestures and looking behaviors between duo pianists emerge to consolidate timing, phrasing, and musical style, to coordinate performance and communicate musical ideas. As to coordination in ensemble performance, Keller (2008) highlighted the cognitive processes of anticipating one's and others' sounds by auditory imagery, dividing attention between own's and others' action, and adjusting one's movements, while Glowinski, Dardard, Gnecco, and Camurri (2014) used head motion tracking to show how musicians in a string quartet pay attention to other performers' head movements to predict their upcoming action, and to maintain group cohesion. King and Ginsborg (2011) found that pianists use more gestures and glances when in duo with their familiar and same-expertise partners than with unknown musicians.

A fourth area of investigation is gestures in singing: in performances of Annie Lennox and Robbie Williams, Davidson (2001, 2006) stressed how performers develop a vocabulary of gestures that accompany their singing, singling out their different functions—to communicate expressive intention, need for coordination with co-performers or audience participation, to signal extra-musical concern, display singer's personality, or show-off—and underscoring the role of gesture in social interaction and in the very construction of the performance.

Within a fifth research area, musical teaching, Simones, Schroeder, and Rodger (2015a, 2015b) found differences in performance by same and different grade piano students depending on the type and frequency of gestures used by their teacher during lessons. Liao and Davidson (2007) evidenced the link between children's singing voice and their use of gesture by observing the Dalcrozian mirror and follow games. Rahaim (2012) stresses the intertwining of gesture and voice in Hindustani music and in teacher–student relationship.

The conductor's body

In the musicological domain, among the conductor's movements, it is mainly gestures that have been studied, and especially concerning their rhythmical import (Green, 1997; Rudolf, 1995; Saito, 1988). In psychological research, various topics have been tackled. Within works on temporal gestures, Luck (2000) studied the synchronization between the movement of the conductor's wrist and elbow and musicians' attack, finding that attack delay is shorter when seeing the baton only than in a full-cue condition; Luck and Nte (2008), through a computer-based environment for manipulating and presenting conductors' gestures, showed that the synchronization ability of musicians is determined by their own previous experience, not by conductors' experience or by the radius of curvature of the beat. Parton and Edwards (2009), by coding conductors' gestures in terms of size, handshape, body orientation, position, arm extension, accompaniment by left-hand movement, and measuring musicians' response by the lag from the completion of the gesture stroke to the beginning of audible sound, found that conductors' gestures are systematically related to musicians' action.

Some studies deal with musicians' perception and comprehension of conductors' body movements. Using a spatial occlusion approach, Wöllner (2008) demonstrated that viewing the conductors' face only is more effective in understanding expressive meaning, whereas viewing arms only is better for amount of information. The importance of facial expression also in order to conductor gestures comprehension is demonstrated by Mayne (1992).

Several works investigate the effects of conductors' nonverbal behaviors on the produced music (Gritten and King, 2011). By comparing more and less "musical" band performances, Grechesky (1986) found that the conductor's communication determining better performance includes more body movement, more use of left hand, coordination of both hands, use of verbal imagery, emblems, illustrators, iconic behaviors, and approving facial expressions. Yet, higher expressivity does not always result in more effective conducting. Price and Winters (1991), by comparing strict and expressive conducting behaviors, show that eighth-grade band performers have a better opinion of more expressive conductors who frequently use body movements, expressive gestures, eye contact, and facial expressions of approval and disproval, even though their performance does not significantly differ from one of bands conducted by a strict conducting style. By testing the use of hand and arm gestures, facial expression, eye contact, and body movements of expert and novice conductors, Byo and Austin (1994) found that expert conductors, all equally successful, achieve their effectiveness through divergent expressive styles that differ from those of novice conductors. Van Weelden (2002), tackling the physical appearance of conductors, the ways it fits their stereotypical role, and how it affects opinions on their performance, did not find a strong relationship between eye contact and performance ratings, but did find a moderate relationship between conductors' facial expression, evaluators' confidence in the conductor, and general conducting effectiveness. A broader perspective is taken by Dineen (2011), who compared how ice hockey coaches and conductors deal with the economy of gesture, the mean expenditure needed to achieve an outcome, and starting from the downbeat that marks the first note of a performance overviews several aspects that make an effective conductor: from the semiotic competence—the ability to convey the meaning of a gesture clearly—to the ability to be followed by the ensemble even when a shadow ensemble is present (a group of musicians assuming the leadership).

Other studies focus on gestural and facial behaviors from the point of view of their meaning. Boyes Bräm and Bräm (2004) and Veronesi (2009) focus on the iconicity of conducting gestures, and their underlying metaphors, and Davidson (2001) refers to a "vocabulary" of singers' or conductors' gestures. Poggi (2011) stresses that the conductor's communication

is multifunctional because it takes care of all the parameters of music (melody, rhythm, harmony, intensity, expression) and not only its technical but also its aesthetic aspects. In both concert and rehearsal, it provides information about the sound to perform: who should sing or play, when, how; what semantic contents to express by words and music, what melody, rhythm, tempo, timbre, intensity, expression, what musical structure to produce; and it does so multimodally, by hands, eyes, trunk, face, the whole body. All of these signals constitute a proper lexicon, that is, a relatively steady and shared list of signal-meaning pairs (Poggi, 2007). After the lexicon of face, Poggi (2002, 2018) and Poggi and Ansani (2018) analyzed the conductor's lexicon of gaze, showing that it conveys a wide range of meanings, from concentration to attack, from intensity to emotions, from requests for accuracy to praise and reproach.

In the domain of Human–Computer Interaction, analyses of conductors' gestures are the basis for Bos, Reidsma, Ruttkay, and Nijholt (2006); Castellano, Bresin, Camurri, and Volpe (2007); and Johannsen and Nakra (2010) to build conducting interfaces.

In order to analyze performers' or conductors' movements, many of the above works exploit coding schemes that, in correspondence to specific bars of the performance, describe them in terms of Ekman and Friesen's Facial Action Coding System for facial expression, or of physical parameters of body movement, such as "elbows circle out," "head nod," and "side-to-side sway" (Davidson, 2012), and classify them in terms of Ekman and Friesen's (1969; 1978) or other typologies of nonverbal behaviors; yet only some of them attribute specific meanings to the described movements, and finally single out a vocabulary of performers' (Clarke & Davidson, 1998; Davidson, 2001) or conductors' signals (Poggi, 2006, 2017, 2018).

The conductor's communication systems

This work is part of a broader project aimed at finding out the repertoire of the conductor's body communicative signals.

We start from a cognitive model of communication (Poggi, 2007), according to which a communicative process occurs anytime a Sender S has a goal for an Addressee A to come to know a new Belief B, and to this end produces a communicative Signal *s*, a perceivable stimulus that is linked to Belief B, which is then its meaning, through a Communicative System CS, a system of rules to set correspondences between signals and meanings, represented in both Sender's and Addressee's minds. The goal of communicating need not be a conscious one, like for *saying hello* or *waving a hand*, but may also be a goal S is not aware of, like *raising eyebrows* to emphasize a word, or even a biological function, such as *blushing* for shame. Communicative signals may be distinguished in terms of various dimensions: their perceptual modality (vision, audition, olfaction, taste, touch), their productive modality (words, prosody, gestures, facial expressions, gaze, posture), and the type of information they convey (concerning the world, for example, when communicating how to shape the mouth when singing a high note; concerning the Sender's identity, for example, when a gesture expresses the singer's personality; and concerning the Sender's mind, for example, the facial expression of an emotion). Another dimension is the "cognitive construction" of signals: some are "codified," that is, steadily represented in Senders' and Addressees' memory thus forming a "lexicon," that is, a list of signal-meaning correspondences. A typical case of codified signals are words of languages, and symbolic gestures, those that have a conventional shared meaning in a given culture. To produce and understand codified signals, one must simply access the lexical entry in memory, so they can be totally arbitrary signals, with no link or resemblance with their meanings: for example, *index finger perpendicular to closed Mouth* meaning "silence." Other signals are "creative," brand new signals

invented on the spot: like a neologism the first time it is used. To allow an Addressee to understand such signals, their meaning must be easy to infer; so "creative" signals cannot be totally arbitrary but mostly iconic, that is, their physical appearance imitates, resembles some perceivable aspect of the meaning, like some shape, size, location, or movement: for example, a conductor may ask for a *pianissiMo* by *squinting eyes*, since smaller eyes "resemble" smaller sound. These signals, being invented on the spot, do not form a lexicon, that is, a memorized list of *correspondence rules*, but are nonetheless constructed following some *inference rules* stating that the signal imitates perceivable aspects of the meaning, and which ones. After its invention, the iconic signal may become codified in its turn, that is, get steadily represented in memory and become part of a lexicon. Signals of all modalities exploit these two cognitive construction devices, one implying memorized correspondence rules, the other inference rules. Investigating the communication system of conductors' signals implies to single out both kinds of rules: on one hand, to write down a lexicon of their codified signals, on the other to single out the construction rules of their creative ones.

Signals of intensity

Within the long-lasting project of outlining repertoires of the conductors' signals, to accomplish subsets of this endeavor, two routes can be followed. One is to focus on a single modality and to search for all the signals used by conductors in that modality. This has been done, for instance, for the conductor's lexicon of gaze (Poggi, 2018; Poggi & Ansani, 2018; Poggi et al., 2020), where all the gaze items used in conduction were listed, whether conveying emotions, indications of intensity, pitch, or expressivity. Another route is to select a semantic area, among the meanings that conductors must convey, and to search for the signals of either one or all the modalities that provide indications in this area: for example, one could investigate the gestures of attack or closing. Here we choose this approach focusing on a specific semantic area, the body signals conveying indications of intensity—*piano*, *forte*, *crescendo*, *diMinuendo*—and present two studies: in a coding study (see Poggi, 2017; Poggi & Ansani, 2016, for more details), we analyzed video-recorded fragments of orchestra and choir performance and singled out the conductors' gestures, postures, and gaze items that convey intensity indications; then we conducted a perception study on a small number of gestures of intensity. Here we do not deal with "musical gestures" in Hatten's (2006) sense, but with the "physical gestures" used by the conductor to convey indications for performance. Following our model, we define as gesture any movement produced by hands, arms, and shoulders with the goal of conveying some meaning: an emotion, a conceptual content, a mental image.

A coding study on intensity body signals

To single out the signals of intensity used in conducting, we first ran a qualitative observational study: our hypothesis was that the communicative movements of the conductor are recurrent and systematic, that their meanings are shared between maestro and musicians, that a subset of them is devoted to conveying intensity indications, and that they can be singled out writing down a repertoire of them.

Corpus and analysis

Ten fragments were collected of orchestra and choir conducting by three conductors: two from rehearsals by Riccardo Muti and Leonard Bernstein, and eight (four from concerts, four from

rehearsals) by Alessandro Anniballi, the Conductor of choir "Orazio Vecchi," an amateur choir in Rome, for a total length of 122'35": 27'55" of concert and 94'40" of rehearsal. The fragments were selected based on the clear visibility of the conductor's body, and they were analyzed through the annotation scheme in Table 1.

In this annotation scheme, the conductor's body signals soliciting dynamic action are analyzed in detail: column 1 contains the time in the video, column 2 the dynamic indication written on the musical score, column 3 (in case of choir performance), the words sung at the same time of the conductor's signal analyzed; column 4, the modality under analysis (posture, gesture, head, face, gaze); column 5, a description of the analyzed signal in terms of its parameters, that is, body parts involved and movements: for instance, in Table 1, the conductor's posture is described as "Trunk forward, shoulders closed, head forward downward"; column 6 contains the goal of the body movement performed, which counts as its "originary meaning"—that is, the goal of the "bodily" meaning, from which the communicative meaning stems (Poggi, 2007): in this example, for instance, the posture is that of someone making himself smaller. Column 7 contains the meaning attributed to the movement described in column 5: in this case, a posture of bending to get smaller means "play/sing softer," that is, "make a 'smaller' sound." Column 8 classifies the signal according to the typology illustrated below, and column 9 clarifies its underlying semiotic device. For instance, making oneself smaller is an iconic posture (column 8) that exploits a transmodal shift (column 9), a metaphoric transfer from space to sound: a smaller body evokes a smaller (i.e., softer) music; a body that takes less room recalls a sound that takes less energy.

What distinguishes this annotation scheme from others in literature is column 7, which attributes a meaning to each signal. According to our model, in fact, each communicative movement by definition bears a meaning that can be rephrased in a verbal language. Moreover, since the conductor's communication is multimodal, the synchrony of signals in all modalities is represented by this annotation. In some cases, the same meaning of intensity is conveyed by two or more modalities at the same time: in Table 1, the *bending posture* means "softer," the gesture of *folded elbows* means "quiet," and *raised eyebrows* means "light"; but some modalities convey different—yet not contrasting—meanings: *head forward* means "pay attention."

How are the signals classified in columns 8 and 9? Based on the correspondence between the signal in column 5 and the meaning in column 7, the following set of semiotic devices was found to govern the construction of all signals of intensity:

1. *Generic syMbolic signal*: a signal also used by laypeople (not expert musicians): an emblem (Ekman & Friesen, 1969) or other culturally codified signal (Poggi, 2007). For example, *Index finger over lips*, a symbolic gesture meaning "be silent," is often used by conductors to ask for "piano";
2. *Specific syMbolic signal*: a codified signal used in conducting with a more specific or a slightly different meaning that one used in everyday life: for example, *hands, palM up, oscillating on wrist up-down* in its everyday use (Morris, 1977; Poggi, 2007) means "come here" since it encourages the Addressee to come closer to the Sender. In conducting it means "come on, play/sing louder";
3. *Direct iconic signal*: a gesture imitating some movement or another kind of change in another modality (Poggi, 2008). The conductor's *curve arMs widening*, imitating a swelling body, is an iconic gesture asking for a *crescendo*: a swelling sound.
4. *Indirect iconic signal*: a gesture or other signal that does not directly imitate the movement it refers to, or its transmodal analogue, but some movement that by inference may recall the desired intensity. Such indirect iconicity passes through two kinds of movement:

Table 1. An annotation scheme for signals of intensity.

1	2	3	4	5	6	7	8	9
Time	Score	Words	Modality	Signal description	Originary meaning	Meaning	Signal type	Semiotic device
3.18	*Più piano*	Je-e-su-u Chri-i-**i-ste**	Posture	*Trunk forward, shoulders closed, head forward downward*	I make myself smaller >	Softer	Iconic	Transmodal iconicity: space > sound Take less room = make softer sound
			Gesture	*Elbows folded*		please quiet		
				Rh. Open, palM down, pats downward		attenuate	Symbolic gesture	Generic Codified
				Lh. palM forward, Move backward	*refrain froM* >			
						be careful: do something	Iconic spatiographic gesture	Transmodal space > sound
				Lh. open spread fingers perp. to palM, Makes precision grip	I pick up something carefully	subtle	Iconic spatiographic gesture	Take less room = make softer sound
								Transmodal Movement > sound Smaller grip = softer sound
			Face	*Head forward*,		I ask for attention		
				tense lips		Be careful		
			Gaze	*Eyebrows raised* *Eyelids tense, squeezed half-closed*	I try not to be heavy	Be light,	Indirect iconic	Help movement

Table 2. The conductors' signals of intensity.

	Forte	Piano	Crescendo	Diminuendo	Total
Gesture	33	23	5	2	63
Head	3			1	4
Face	5	1			6
Gaze	7	7		1	15
Posture		3		1	4
Body	2		3	1	6
Voice		1		1	2
Total	50	35	8	7	100

4.a. *Motor attitude*: the signal imitates a movement that is usually performed when producing another movement or the resulting sound: for example, to mean *sforzato* (with effort) the conductor suddenly *squints his eyes*, imitating the movement people do when striving to perform some physical action;

4.b. *EMotion expression*: the signal imitates the movements typically performed in the expression of an emotion that, when felt, induces the desired type of attitude or movement: *hands in the shape of claws vibrating with high Muscular tension* request a "forte," because tension is typical of an activating emotion like anger, and anger mobilizes the energy required for playing or singing *forte* (Poggi, 2006).

Results

From this qualitative analysis, a repertoire of 100 signals of intensity was singled out—gestures, head movements, facial expressions, gaze items, postures: 50 for *forte*, 8 for *crescendo*, 35 for *piano*, and 7 for *diMinuendo*, distributed as shown in Table 2.

Within these 100 intensity signals, 63 are gestures, more precisely *gesture tokens*, that is, single actual occurrences of gestures, but since, in 21 cases out of the 63, different occurrences have the same handshape, location, and movement corresponding to the same meaning, 21 *gesture types* emerge in total, 8 requesting "forte," 7 "piano," 4 "crescendo," and 2 "diminuendo."

Out of the 21 items, 5 are symbolic gestures (e.g., *Both hands open with close fingers, palMs down, slightly push downward*, meaning "keep down, keep calm," conveys "piano") and 16 are iconic gestures, representing objects (*both hands open with close fingers, palMs down, Moving outward* as if smoothing a flat surface, meaning "piano") or actions (*arMs with open hands curve opening outward with shoulders raising upward* = "swelling," a gesture for *crescendo*).

Any gesture can be analyzed in terms of the values it assumes on a set of parameters: handshape, location, movement, and orientation (Stokoe, 1978; Volterra, 1987), and within movement the sub-parameters of expressivity: amplitude, fluidity (Hartmann, Mancini, & Pelachaud, 2002), and tension (Poggi, 2007). For example, as to handshape a gesture may assume the values (the specific handshapes) *fist, extended index finger, open hand*, or other.

Therefore, within the analyzed items, it is not only a gesture as a whole that can count as dynamic indication, but even a single value in one parameter: its handshape, movement amplitude, or duration. For instance, the *fist* handshape by itself conveys a metaphor (Boyes Bräm, 1981) of strength, hence it asks for "forte"; as to the "expressivity parameters" (Hartmann, Mancini, & Pelachaud, 2002; Poggi & Pelachaud, 2008) of tension, fluidity, amplitude, for example, an *increasing aMplification* of gesture or body size means "crescendo."

"Forte" is typically conveyed by the handshapes *fist* and *claw*, direction *toward performers, and jerky, wide, high tension* movements, all metaphorically expressing an idea of strength and energy. "Piano" is evoked by soft and fluid movements, downward direction, low tension, amplitude, and quantity of movement.

A perception study on intensity gestures

After putting forward our hypotheses about the meanings of intensity gestures by our observational study, we designed a perception study to test them.

Research questions and working hypotheses

Our main research question was, how shared is the repertoire of musical intensity gestures across people? Do people interpret these gestures the same way as hypothesized above? What is the basis for their interpretation, that is, what do people precisely look at, in a conductor, to interpret intensity signals? Finally, are these interpretations the same for experts, that is, people acquainted with music performance, and for non-experts, that is, people who have no experience in playing or singing in a musical ensemble? To answer these questions, we selected eight gestures out of the ones found in the coding study to test their interpretation by experts and non-experts in music.

Our first hypothesis was that these gestures were attributed the same meanings as posited from our analysis. Our second hypothesis was that to interpret each gesture, expert participants would look more at the conductor's hands than at other parts of his body. Our third hypothesis was that such interpretation was easier or more frequent for expert musicians than for non-experts in music.

Method

Experimental design

The quasi-experimental design of our perception study was bifactorial 4 × 2 between subjects: the independent variables were gesture meaning (piano/forte/crescendo/diminuendo) and musical expertise (expert/non-expert); the dependent variables were as follows:

1. perception of the intensity level requested by a conductor's gesture, measured on a 5-point Likert-type scale;
2. open answers on gesture meaning;
3. participant's self-reported gaze direction, respectively, onto the conductor's Hands, Face, Body (as a whole), Gaze, Mouth, Head, and Trunk, each measured on a 5-point Likert-type scale.

Participants

Seventy-seven participants (45 females; mean age: 39, $SD = 19.47$) participated in the study: 44 non-experts (28 females; mean age = 28.3) and 33 experts (17 females; mean age = 51.6).

The majority of non-experts were recruited on a voluntary basis among students in an introductory psychology course, while to increase the chance to find experts in music, we recruited

a fair number of participants among singers in amateur choirs and musicians playing in amateur ensembles, whose level of expertise was double-checked by means of specific questions at the end of the experiment (see "Procedure" section).

Materials

Our stimuli were eight brief videos (1–4 min, median length = 2 min) of two conductors, Herbert von Karajan and Alessandro Anniballi, each performing four gestures of the same types as those found in the coding study, respectively, with the following four meanings: piano, forte, crescendo, and diminuendo. Each participant was shown both the well-known and the amateur conductor, to check for differences in interpretation accuracy.

Procedure

Our challenge was to see whether participants could figure out the meaning of each gesture without listening to the music and even by themselves; but if this was not the case, we wanted to obtain some answer anyway. So, after answering some socio-demographic questions and with no anticipation about the topic of the questionnaire, each participant was asked to watch two (randomly assigned) gestures (without audio), and to guess the meaning of the first by an open question, and of the second by a Likert-type scale question: the former aimed at leaving participants free to figure out a meaning by themselves, the latter to narrow their guess about the second gesture to only a few alternatives.

The open question on the first video was, "In this video, the conductor is asking the orchestra to play in a certain way. By a sentence express what he is communicating in your opinion," followed by a further open question: "How can you tell this is so?" to understand the bases for the participant's interpretation.

The Likert-type scale question concerning the second video was, "In this video the conductor is asking the orchestra to play in a certain way. Check how plausible you think is each of the following (from 1 = *not at all* to 5 = *very Much so*)," with the alternatives being eight possible randomized meanings: four of intensity ("play soft," "play loud," "play progressively louder," and "play progressively softer"), corresponding in everyday language to *piano, forte, crescendo*, and *diMinuendo*, plus four distractors: "play more in tune," "play in a passionate way," "scan notes well," and "keep the time."

After the closed question, participants were asked what part of the conductor's body had they mainly looked at to understand the meaning: hand movements, head movements, facial expression, gaze, mouth movements, trunk position, or body movement.

The two videos were always one with Karajan's and one with Anniballi's gesture, in random order. Before the first video, the participant answered questions about socio-demographic information and, to achieve a more precise assessment of the level of music expertise, they were asked whether they used to play or sing in a music ensemble, if so, how often, what music genre she or he liked most (classical, folk, rock, jazz, songwriting), and how frequently she or he attended concerts.

After viewing the first stimulus and verbally phrasing its meaning, participants were asked some questions about music theory (types of instruments, voice timbres), for two reasons: first to test their self-report about musical competence, second to distract them before the second stimulus question. In such a way, the Likert-type question did not immediately follow the open one, in order to prevent participants' second answer from being affected by their first answer.

Table 3. ANOVA—Interpretation of intensity gestures.

Perceptual evaluation (item)	Stimulus	*n* of views	Mean	*SD*	$p < .001$	
					df	*F*
Suonate piano (play soft)	Piano	20	3.4	1.69	3	12.76
	Forte	14	1.00_a	0		
	Crescendo	23	2.00_a	1.27		
	Diminuendo	20	3.45	1.53		
	Total	77	2.55	1.65		
Suonate forte (play loud)	Piano	20	1.75_c	1.33		
	Forte	14	3.85	1.16	3	9.26
	Crescendo	23	2.65	1.66		
	Diminuendo	20	1.55_c	1.23		
	Total	77	2.35	1.6		
Suonate progressivamente sempre più forte (play progressively louder)	Piano	20	1.90_d	1.44		
	Forte	14	3.57	1.39		
	Crescendo	23	3.73	1.51	3	10.56
	Diminuendo	20	1.85_d	1.22		
	Total	77	2.74	1.64		
Suonate progressivamente sempre più piano (play progressively softer)	Piano	20	2.65	1.63		
	Forte	14	1.35_b	0.84		
	Crescendo	23	1.65_b	1.22		
	Diminuendo	20	3.3	1.59	3	7.67
	Total	77	2.28	1.56		

ANOVA: analysis of variance; *SD*: standard deviation.
Subscripts a, b, c, and d show the results of the post hoc Tukey test for Intensity with significance level at $p < .05$; a: *Piano* versus other levels: b: *Diminuendo* versus other levels; c: *Forte* versus other levels; d: *Crescendo* versus other levels.

Data analysis

All quantitative data were analyzed using SPSS 26.0 with analyses of variance (ANOVAs) for the continuous variables such as the intensity gesture recognition score, having as independent variables the type of gesture (*piano, forte, crescendo, diMinuendo*) and the ex-post calculated level of expertise (expert vs. non-expert, according to the mean score). A qualitative analysis was run for answers to the open question on the gestures' meaning; these answers were then treated as nominal variables analyzed with a non-parametric test (chi-square).

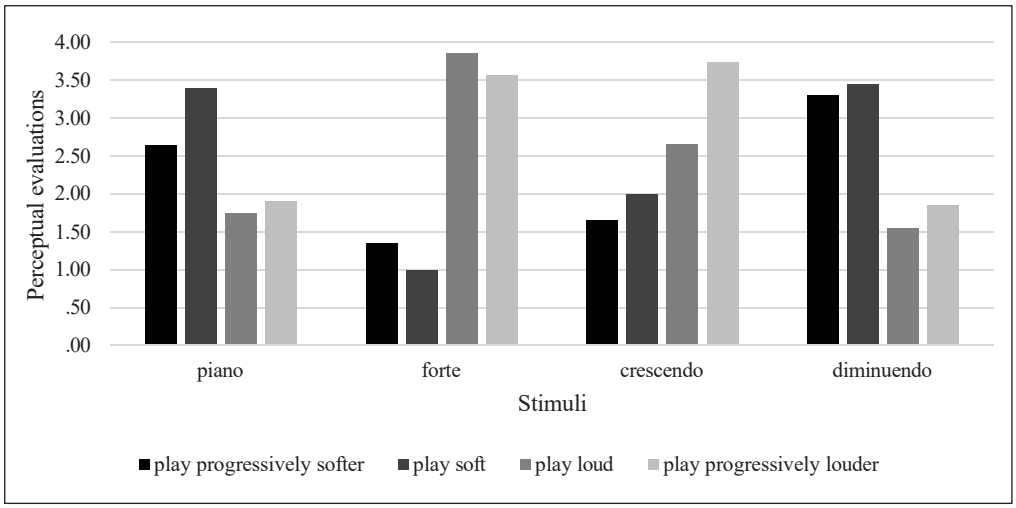
Results

Our manipulation (see ANOVA in Table 3) was successful in that the stimuli produced statistically significant differences on the items ($p < .001$).

Multiple choice interpretation of intensity gestures

Our main question was whether participants would attribute to the stimulus gestures the same meaning of intensity hypothesized by our coding study (Table 2). The ANOVA analysis (Table 3) pointed out a significant effect of the independent variable gesture intensity (*piano*/*forte*/*crescendo*/*diMinuendo*) on each item: the stimulus gesture was perceived as conveying

Figure 1. Interpretation of Intensity Gestures.

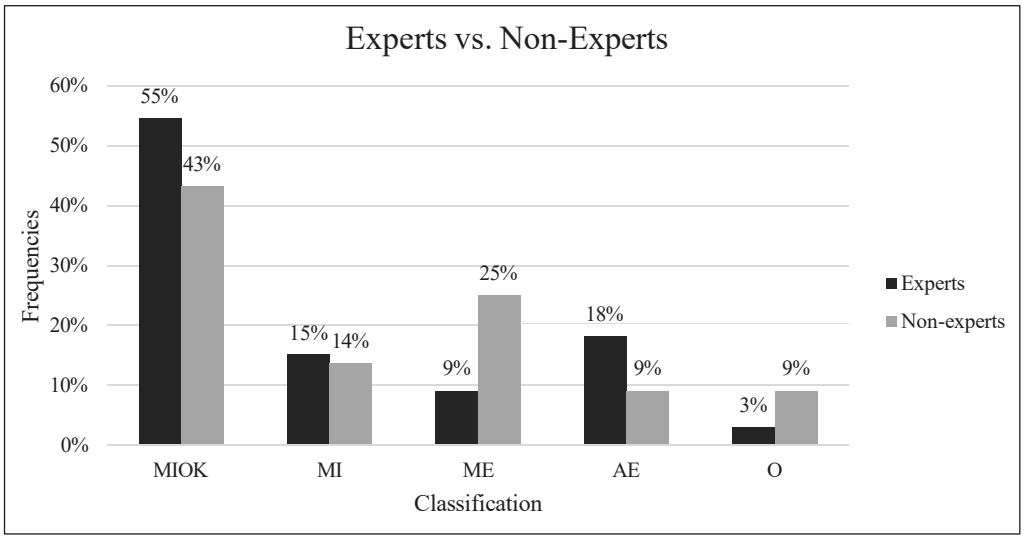


the same intensity as hypothesized in three cases out of four: the gesture *piano* mainly elicited the answer "play soft" (3.40), *forte* mainly elicited "play loud" (3.86), *crescendo* elicited "play progressively louder" (3.74), and *diMinuendo* mainly elicited "play soft" (3.45) (Figure 1). A Tukey post hoc test (Table 3 subscripts) revealed that interpretation followed the polarity of the gesture: the significance of the item *piano* is given by the difference between *piano* and *diMinuendo* gestures from *forte* and *crescendo* ones; the significance of the item "play loud" is given by the mean differences between *forte* and *crescendo* from *piano* and *diMinuendo*; the significance of the item "play progressively louder" is given by the difference of the gestures *forte* and *crescendo* from *piano* and *diMinuendo*, and finally the significance of the item "play progressively softer" is given by the difference between *piano* and *diMinuendo* from *forte* and *crescendo*. The gestures for lower intensity, that is, for *piano* but even more for *diMinuendo*, are more frequently interpreted as requesting low intensity, namely, as "play soft" and "play progressively softer"; and symmetrically the gestures for *forte* and *crescendo* are more often understood as "play loud" and "play progressively louder." Yet, on the same end of the intensity continuum, participants do not discriminate (ratings are quite similar) between *piano* and *diMinuendo* on the one side and *forte* and *crescendo* on the other. As a matter of fact, we found a correlation between *piano* and *diMinuendo,* $r = .51$, $p < .01$, and between *forte* and *crescendo,* $r = .60$, $p < .01$.

We may account for this result by recalling that the differentiation between *piano* and *diMinuendo* or between *forte* and *crescendo* is more a feature of the musical score that involves a sort of linearity within the internal music flow than a perfectly conveyable and discernible indication. In other words, the meaning conveyed by the conductor in both cases simply means something like "play softer/play louder than you are currently doing", regardless of any "time linearity."

An interesting detail emerges from these data concerning the relationship between intensity and emotion. As mentioned above regarding indirectly iconic intensity gestures, the expression of emotions is one of the strategies used by conductors to convey dynamic indications. Here we observed some correlations in this direction in participants' interpretations across all videos: the item "play in a passionate way" (formally, a distractor) showed a strong correlation with both the stimuli *forte,* $r = .60$, $p < .01$, and *crescendo,* $r = .52$, $p < .01$. We may account for this

Figure 2. Participants' Free Interpretation of Intensity Gestures.



result by supposing that our participants often understood the right dynamic indication only passing through the request for passion performed by the conductor.

Free interpretation

Concerning the open question on the first stimulus, by qualitative analysis of the participants' answers, we classified them into five categories:

- MIOk (Musical Intensity Ok): participant correctly interpreting the gesture, for example, stimulus *forte* interpreted as "intensify the sound," "play louder"
- MI (Musical Intensity): participant correctly understanding that the meaning was about intensity, but not guessing the right one, for example, "vocal diminuendo" (instead of a simple "piano")
- ME (Music—Else): participant understanding a meaning linked to musical features other than intensity (pitch, rhythm, tempo, etc.), for example, "play faster," "play using a higher pitch" (instead of "play louder")
- AE (Attitude—Expressivity): participant understanding a meaning linked to feelings and their relative expressions, for example, "majesty," "supplication," "tenderness"
- O (Other): participant answering something irrelevant or hard to interpret, for example, "look at me"

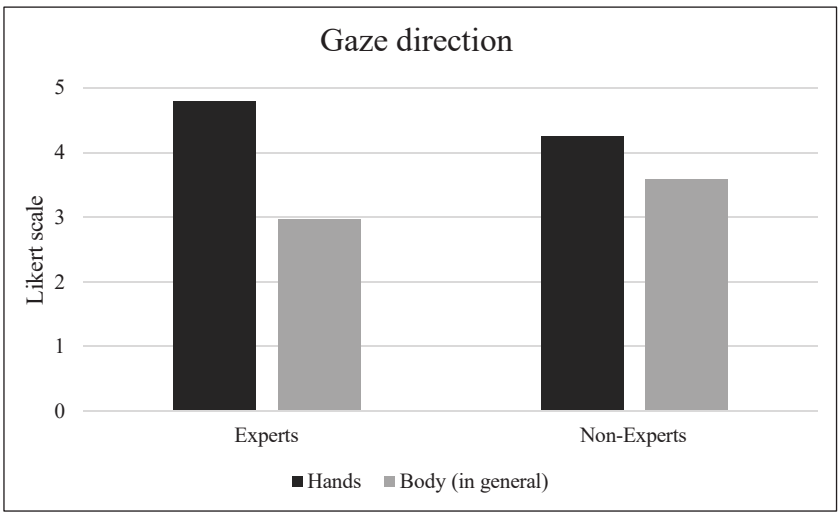
The five categories are ordered according to their correctness, that is, their closeness to the correct meaning of each stimulus based on the previous qualitative study. The trend in Figure 2 might answer our question. The relatively higher frequency of MIOk answers (participant guessed not only a musical and an intensity indication, but the right one) seems to show that the stimuli are quite easily correctly recognized, and not only by music experts. On the contrary, less pertinent answers (classified as Music—Else and as Other) look more frequent in non-experts as opposed to experts.

Table 4. Body parts looked at to interpret the intensity signal.

$n=77$	Mean	*SD*
Hands	4.48	0.95
Facial expression	3.9	1.28
Body (in general)	3.32	1.40
Gaze	2.92	1.43
Mouth	2.76	1.50
Head movements	2.64	1.34
Trunk	2.57	1.37

SD: standard deviation.

Figure 3. What Do Experts and Non-Experts Look At.



What are you looking at?

To assess what in a gesture most clearly bears the meaning of intensity, for each video, after asking to interpret the gesture, we asked participants what they were looking at to understand that meaning, by seven items with 5-point Likert-type scales: hands, head movements, facial expression, gaze, mouth, trunk, body (in general).

In general, the hands are the body part most frequently considered to understand intensity indications (Table 4). This is a robust result because it does not significantly vary for either expert or non-expert participants. Nevertheless, from an ANOVA test looking for different patterns between non-experts ($n = 44$) and experts ($n = 33$), we obtained statistically significant results for hands ($M_{exp} = 4.79$, $M_{non\text{-}exp} = 4.25$), $F(1, 77) = 6.41$, $p = .01$, and body in general ($M_{exp} = 2.97$, $M_{non\text{-}exp} = 3.59$), $F(1, 77) = 3.80$, $p = .05$ (Figure 3): experts focused and relied proportionally more upon the conductor's hands, whereas non-experts tended to a more global evaluation (Figure 3).

Moreover, to determine possible differences depending on which conductor performed the signal (whether Anniballi or Karajan), we checked the conductor as a control variable, but we

did not find any discrepancy, $F(1, 77) = 1.84$, $p = .17$: both experts and non-experts mainly look at the conductor's hands, whoever the conductor may be, to understand what he means.

Experts versus non-experts

The difference between experts and non-experts in their interpretation of gestures was not statistically significant (all ps $> .21$) except for a marginal effect on the item "play forte" ($M_{exp} = 2.06$, $M_{non\text{-}exp} = 2.57$), $F(1, 77) = 3.20$, $p = .7$. We also found two weak interaction effects between the hypothesized meaning of the gesture and the experts/non-experts condition in the item "play loud" and in the item "play progressively louder."

Taken together, the two results suggest that expert participants have a finer-grained interpretation of stimuli since they understand a more technical meaning such as *crescendo* at a higher degree, whereas the non-expert subsample shows some more general interpretations polarizing the answers within the "play loud" item.

Discussion

The two studies presented, a coding and a perception study, investigate the existence of a shared repertoire of body signals in orchestra conductors, focusing on the gestures of intensity. The coding study, through qualitative analysis, singled out 21 gestures and some values in their parameters by which conductors ask musicians to sing or play *forte*, *piano*, *crescendo*, or *diMinuendo*. These gestures exploit the same semiotic devices as everyday communication: symbolic gestures used with the same meaning by laypeople or with a more specific "musical" meaning; and iconic gestures (or values in their parameters) building up their meaning on top of natural motor attitudes or emotional expressions.

A subsequent perception study investigated if eight of those gestures are actually interpreted like hypothesized by the coding study (tackled by an open and a closed question), what parts of the conductor's body participants mainly look at to interpret his gestures, and if music experts and non-experts interpret them differently.

The answer to our first question, whether intensity gestures have systematic and shared meanings, is generally positive: the gestures for the two ends of musical intensity, *piano* and *forte* are fairly distinguished by participants, and confused at most with their shifts of intensity, *diMinuendo* and *crescendo*. As to the second question, as hypothesized, it resulted that the parts of the body looked at to interpret gestures are mainly hands. As to the third question, our perception study unexpectedly did not reveal significant differences between experts and non-experts: musical expertise, at least at a preliminary level, is not strictly needed to understand intensity gestures.

Such easy recognition can be accounted for by an important aspect highlighted by the coding study (Poggi & Ansani, 2016): the high iconicity of intensity gestures. Their direction, speed, tension, and amplitude of movements is highly parallel and isomorphic to parameters of sound and of the motor actions required to produce it.

Some limitations of our perception study could be overcome in future works: samples could be broadened by recruiting a larger sample, balancing the number of experts and non-experts, and participants from different cultures could be included; furthermore, it would be interesting to test possible differences in the production of intensity gestures between orchestra and choir conductors, and between male and female conductors.

Moreover, since our stimuli showed a big portion of the conductor's body, the interpretation of the hand gesture could be affected by the concomitant gaze or head movement. In future, "cleaner" stimuli could be built, for instance, by simulating the gestures to be assessed

in Virtual Agents (Castellano et al., 2007), which afford perhaps less ecological but more controlled stimuli.

Conclusion

This work can be framed within a general project aimed at writing down the communication system of conductors' body signals. Carrying out such an endeavor implies finding out all the signals exploited in each single body modality, as has been done, for example, in studies listing all the gaze items used by conductors (Poggi, 2018), but also focusing on single semantic areas, and to see how the various body modalities convey those meanings. The qualitative analysis of our coding study has singled out, among all intensity signals, 21 intensity gesture types, while the perception study has demonstrated that eight of those gestures are generally correctly interpreted by both music experts and non-experts.

This shows that a repertoire of conductors' signals can be built, but also that this "language" is so iconic and so close to everyday communication that even people not acquainted with ensemble conducting can understand it; nonetheless, this is anyway a technical communication system, covering all the meanings a conductor may need to convey.

Understanding if there are, and what are the correspondences between signals and meanings in the body language of orchestra and choir conductors may contribute to the construction of artificial systems that perform, understand, interpret, or judge music: like ones for home conducting (Friberg, 2005) and conductor virtual agents and robots (Borchers, Lee, Samminger, & Mühlhäuser, 2004; Bos et al., 2006; Friberg, Bresin, & Sundberg, 2006; Ruttkay, Huang, & Eliëns, 2003). On the pedagogical side, studying conductors' signals may enhance both their own and the musicians' awareness of the communicative tools they may rely on during performance, leading to better conductor-ensemble tuning. On the theoretical side, it allows us to compare the cognitive and communicative devices at work in this domain with those holding in other communication systems, including those of everyday communication, to find out their differences and commonalities, to discover more general mechanisms, and to achieve a broader and deeper comprehension of human communication.

Acknowledgements

We are indebted to Valentina Vignoli for her help in the English draft. Our most special thanks to the Choir "Orazio Vecchi," and to Alessandro Anniballi for his fascinating conducting gestures.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is partly funded by the Italian MIUR Project "Cultural Heritage Resources Orienting Multimodal Experiences (CHROME)," grant no. 2015WXBPYK, and partly by a scholarship for research on music, Roma Tre University, grant no. 1236-2019.

ORCID iDs

Isabella Poggi https://orcid.org/0000-0002-4067-3424

Alessandro Ansani https://orcid.org/0000-0002-0657-5732

References

Borchers, J., Lee, E., Samminger, W., & Mühlhäuser, M. (2004). Personal orchestra: A real-time audio/video system for interactive conducting. *MultiMedia Systems*, *9*, 458–465.

Bos, P., Reidsma, D., Ruttkay, Z., & Nijholt, A. (2006). Interacting with a virtual conductor. In R. Harper, M. Rauterberg & M. Combetto (Eds.), *Proceedings of the fifth international conference on entertainment computing* (pp. 25–30). Berlin, Germany: Springer.

Boyes Bräm, P. (1981). Significant features of the handshape in American Sign Language (Unpublished doctoral dissertation). University of California, Berkeley.

Boyes Bräm, P., & Bräm, T. (2004). Expressive gestures used by classical orchestra conductors. In C. Mueller & R. Posner (Eds.), *Proceedings of the symposium on the semantics and pragmatics of everyday gestures* (pp. 127–143). Berlin, Germany: Weidler.

Byo, J. L., & Austin, K. R. (1994). Comparison of expert and novice conductors: An approach to the analysis of nonverbal behaviors. *Journal of Band Research*, *30*, 11–34.

Castellano, G., Bresin, R., Camurri, A., & Volpe, G. (2007). Expressive control of music and visual media by full-body movement. In *Proceedings of the 7th International conference on new interfaces for musical expression* (pp. 390–391). New York, NY: Association for Computing Machinery. Retrieved from https://dl.acm.org/doi/abs/10.1145/1279740.1279829

Clarke, E. C., & Davidson, J. W. (1998). The body in performance. In W. Thomas (Ed.), *Composition, performance, reception: Studies in the creative processes of music* (pp. 74–92). Aldershot, UK: Ashgate.

Davidson, J. W. (1991). The perception of expressive movement in music performance (Unpublished doctoral dissertation). City, University of London, London, England.

Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, *21*, 103–113.

Davidson, J. W. (1994). Which areas of a pianist's body convey information about expressive intention to an audience? *Journal of Human Movement Studies*, *26*, 279–301.

Davidson, J. W. (2001). The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. *Musicae Scientiae*, *5*, 235–256.

Davidson, J. W. (2006). She's the One': Multiple functions of body movement in a stage performance by Robbie Williams. In A. Gritten & E. King (Eds.), *Music and gesture* (pp. 208–226). Aldershot, UK: Ashgate.

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.

Davidson, J. W. (2012). Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. *Psychology of Music*, *40*, 595–633.

Davidson, J. W., & Correia, J. S. (2002). Body movement. In R. Parncutt & G. E. McPherson (Eds.), *The science and psychology of music performance* (pp. 237–250). Oxford, UK: Oxford University Press.

Demos, A. P., Chaffin, R., & Kant, V. (2014). Toward a dynamical theory of body movement in musical performance. *Frontiers in Psychology*, *5*, Article 477.

Dineen, P. M. (2011). Gestural economies in conducting. In A. Gritten & E. King (Eds.), *New perspectives on music and gesture* (pp. 131–158). Aldershot, UK: Ashgate.

Ekman, P., & Friesen, W. V. (1969). The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica*, *1*, 49–98.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA: Consulting Psychologists Press.

Friberg, A. (2005). *Home conducting: Control the overall musical expression with gestures*. San Francisco, CA: International Computer Music Association.

Friberg, A., Bresin, R., & Sundberg, J. (2006). Overview of the KTH rule system for musical performance [Special issue]. *Advances in Cognitive Psychology*, *2*, 145–161.

Glowinski, D., Dardard, F., Gnecco, G., & Camurri, A. (2014). Expressive non-verbal interaction in a string quartet: An analysis through head movements. *Journal of Multimodal User Interfaces*, *9*, 55–68.

Grechesky, R. N. (1986). An analysis of non-verbal and verbal conducting behaviors and their relationship to expressive musical performance (aesthetic, gestural) (Doctoral thesis). The University of Wisconsin–Madison, Madison.

Green, E. (1997). *The modern conductor: A college text on conducting based on the technical principles of Nicolai Malko as set forth in his* The conductor and his baton (6th ed.). Englewood Cliffs, NJ: Prentice Hall.

Gritten, A., & King, E. (2011). *New perspectives on Music and gesture*. Aldershot, UK: Ashgate.

Hartmann, B., Mancini, M., & Pelachaud, C. (2002, June 21). *ForMational paraMeters and adaptive prototype instantiation for MPEG-4 coMpliant gesture synthesis*. Paper presented at the Proceedings of Computer Animation, Geneva, Switzerland.

Hatten, R. S. (2006). A theory of musical gesture and its application to Schubert and Beethoven. In A. Gritten & E. King (Eds.), *Music and gesture* (pp. 1–23). Aldershot, UK: Ashgate.

Johannsen, G., & Nakra, T. M. (2010). Conductors' gestures and their mapping to sound synthesis. In Godøy, R. I., & Leman, M. (Eds.), Musical gestures: Sound, movement, and meaning (pp. 264–298). New York, NY: Routledge.

Juslin, P. N. (2005). From mimesis to catharsis: Expression, perception, and induction of emotion in music. In D. Miell, R. MacDonald & D. J. Hargreaves (Eds.), *Musical coMMunication* (pp. 85–115). Oxford, UK: Oxford University Press.

Juslin, P. N., & Timmers, R. (2010). Expression and communication of emotion in music performance. In P. N. Juslin & J. A. Sloboda (Eds.), Handbook of music and emotion: Theory, research, applications (pp. 453–489). Oxford, UK: Oxford University Press.

Keller, P. E. (2008). Joint action in music performance. In F. Morganti, A. Carassa & G. Riva (Eds.), *Enacting intersubjectivity: A cognitive and social perspective to the study of interactions* (pp. 205–221). Amsterdam, The Netherlands: IOS Press.

King, E., & Ginsborg, J. (2011). Gestures and glances: Interactions in ensemble rehearsal. In A. Gritten & E. King (Eds.), *New perspectives on Music and gesture* (pp. 177–201). Aldershot, UK: Ashgate.

Liao, M. Y., & Davidson, J. W. (2007). The use of gesture techniques in children's singing. *International Journal of Music Education*, *25*, 82–94.

Luck, G. (2000). Synchronising a motor response with a visual event: The perception of temporal information in a conductor's gestures. In *Proceedings of the 6th international conference on Music perception and cognition (ICMPC6)* (CDROM). Retrieved from https://www.escom.org/proceedings/ICMPC2000/Mon/Luck.htm

Luck, G., & Nte, S. (2008). An investigation of conductors' temporal gestures and conductor—Musician synchronization, and a first experiment. *Psychology of Music*, *36*, 81–99.

Mayne, R. G. (1992). *An investigation of the use of facial expression in conjunction with Musical conducting gestures and their interpretation by instruMental perforMers* (Doctoral dissertation). The Ohio State University, Columbus.

Morris, D. (1977). *Manwatching*. London, England: Jonathan Cape.

Parton, K., & Edwards, G. (2009). Features of conductor gesture: Towards a framework for analysis in interaction. In C. Stevens, E. Schubert, B. Kruithof, K. Buckley & S. Fazio (Eds.), *Proceedings of the 2nd international conference on Music coMMunication science* (pp. 3–4). Sydney, New South Wales, Australia: University of Western Sydney.

Poggi, I. (2002). The lexicon of the conductor's face. In P. McKevitt, S. O'Nuallàin & C. Mulvihill (Eds.), *Language, vision, and Music: Selected papers froM the 8th international workshop on the cognitive science of natural language processing, Galway, Ireland, 1999* (pp. 271–284). Amsterdam, The Netherlands: John Benjamins.

Poggi, I. (2006). Body and mind in the Pianist's performance. In M. Baroni, M. R. Addessi, R. Caterina & M. Costa (Eds.), *Proceedings of the 9th international conference on Music perception and cognition (ICMPC)* (pp. 1044–1051). Bologna, Italy: University of Bologna.

Poggi, I. (2007). *Mind, hands, face and body: A goal and belief view of MultiModal coMMunication*. Berlin, Germany: Weidler.

Poggi, I. (2008). Iconicity in different types of gesture. *Gesture*, *8*, 45–61. doi:10.1075/gest.8.1.05pog

Poggi, I. (2011). Music and leadership: The Choir Conductor's multimodal communication. In M. Ichino & G. Stam (Eds.), *Integrating Gestures. The interdisciplinary nature of gestures* (pp. 341–353). Amsterdam: John Benjamins.

Poggi, I. (2017). Signals of intensification and attenuation in orchestra and choir conduction. *NorMas*, *7*, 33–47.

Poggi, I. (2018). Lo sguardo del maestro. In C. Corradi (Ed.), *Cultura popolare, religione diffusa, analisi qualitativa: un sociologo italiano a cavallo tra due secoli*. Studi in onore di Roberto Cipriani (pp. 233–252). Perugia: Morlacchi.

Poggi, I., & Ansani, A. (2016, September 29–30). *Gestures of intensity in orchestra and choir conduction*. Paper presented at the Proceedings of the MMSYM (MultiModalSymposium) 2016, Copenhagen, Denmark.

Poggi, I., & Ansani, A. (2018, June 28–30) . *The lexicon of the conductor's gaze*. Paper presented at the Proceedings of MoCo (Movement Computing) 2018, Genoa, Italy.

Poggi, I., & D'Errico, F. (2010, October 29) . *Cognitive Modelling of huMan social signals*. Paper presented at the Proceedings of SSPW 2010: Social Signal Processing Workshop, Firenze, Italy. doi:10.1145/1878116.1878124.

Poggi, I., & Pelachaud, C. (2008). Persuasion and the expressivity of gestures in humans and machines. In I. Wachsmuth, M. Lenzen & G. Knoblich (Eds.), Embodied communication in humans and machines (pp. 391–424). Oxford, UK: Oxford University Press.

Poggi, I., Ranieri, L., Leone, Y., & Ansani, A. (2020). The power of gaze in music: Leonard Bernstein's conducting eyes. *MultiModal Technologies Interaction*, *4*, 20.

Price, H. E., & Winter, S. (1991). Effect of strict and expressive conducting on performances and opinions of eight-grade band students. *Journal of Band Research*, *27*, 30.

Rahaim, M. (2012). *Musicking bodies: Gesture and voice in Hindustani Music*. Middletown, CT: Wesleyan University Press.

Rudolf, M. (1995). *The graMMar of conducting: A coMprehensive guide to baton technique and interpretation* (3rd ed.). New York, NY: Schirmer Books.

Ruttkay, Z., Huang, A., & Eliëns, A. (2003, June 30) . *The conductor: Gestures for eMbodied agents with logic prograMMing*. Paper presented at the Proceedings of the 2nd Hungarian Computer Graphics Conference, Budapest, Hungary.

Saito, H. (1988). *The Saito conducting Method* (F. Torigai, Trans.). Tokyo, Japan: Min-On Concert Association.

Simones, L., Schroeder, F., & Rodger, M. (2015a). Categorizations of physical gesture in piano teaching: A preliminary enquiry. *Psychology of Music*, *43*, 103–121.

Simones, L., Schroeder, F., & Rodger, M. (2015b). Communicating musical knowledge through gesture: Piano teachers' gestural behaviors across different levels of student proficiency. *Psychology of Music*, *43*, 723–735.

Stokoe, W. C. (1978). *Sign language structure: An outline of the coMMunicative systeMs of the AMerican deaf*. Silver Spring: Linstock Press.

Thompson, M. R., & Luck, G. (2011). Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music. *Musicae Scientiae*, *16*, 19–40.

Van Weelden, K. (2002). Relationships between perceptions of conducting effectiveness and ensemble performance. *Journal of Research in Music Education*, *50*, 165–176.

Veronesi, D. (2009, June 4–5). *"The guy directing traffic": Gestures in conducted iMprovised Music between Metaphor and MetonyMy*. Poster presented at RaAM Workshop: Metaphor, Metonymy and Multimodality, Amsterdam, The Netherlands.

Vines, B. W., Krumhansl, C. L., Wanderley, M. M., & Levitin, D. J. (2006). Cross-modal interactions in the perception of musical performance. *Cognition*, *101*, 80–113.

Volterra, V. (1987). *LIS. La Lingua Italiana dei Segni*. Bologna: Il Mulino.

Wanderley, M. M., & Vines, B. W. (2006). Origins and functions of clarinettists' ancillary gestures. In A. Gritten & E. King (Eds.), *New perspectives on Music and gesture* (pp. 165–191). Aldershot, UK: Ashgate.

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.

Williamon, A., & Davidson, J. W. (2002). Exploring co-performer communication. *Musicae Scientiae*, *6*, 53–72.

Wöllner, C. (2008). Which part of the conductor's body conveys most expressive information? A spatial occlusion approach. *Musicae Scientiae*, *12*, 249–272.