



Article

The Power of Gaze in Music. Leonard Bernstein's Conducting Eyes

Isabella Poggi ^{1,*}, Loredana Ranieri ², Ylenia Leone ³ and Alessandro Ansani ^{1,4}

¹ Department of Philosophy, Communication, and Performing Arts, Roma Tre University, 00146 Rome, Italy; alessandro.ansani@uniroma1.it

² National Musical Academy of St. Cecilia, 00196 Rome, Italy; lor.ranieri@stud.uniroma3.it

³ RAI—Radiotelevisione Italiana, 00195 Rome, Italy; ylenia.leone@rai.it

⁴ Department of Psychology, Sapienza University of Rome, 00185 Rome, Italy

* Correspondence: isabella.poggi@uniroma3.it

Received: 6 April 2020; Accepted: 18 May 2020; Published: 20 May 2020

Abstract: The paper argues for the importance and richness of gaze communication during orchestra and choir conduction, and presents three studies on this issue. First, an interview with five choir and orchestra conductors reveals that they are not so deeply aware of the potentialities of gaze to convey indications in music performance. A conductor who was utterly conscious of the importance of gaze communication, however, is Leonard Bernstein, who conducted a performance of Haydn's Symphony No. 88 using his face and gaze only. Therefore, a fragment of this performance is analyzed in an observational study, where a qualitative analysis singles out the items of gaze exploited by Bernstein and their corresponding meanings. Finally, a perception study is presented in which three of these items are submitted to expert, non-expert, and amateur participants. The results show that while the signal for “start” is fairly recognized, the other two, “pay attention” and “crescendo and accelerando” are more difficult to interpret. Furthermore, significant differences in gaze item recognition emerge among participants: experts not only recognize them more, but they also take advantage of viewing the items with audio-visual vs. video-only presentation, while non-experts do not take advantage of audio in their recognition.

Keywords: music performance; conductor; body lexicons; meanings of gaze; interpretation; music experts; amateurs

1. Introduction

If music is made by musicians' souls and bodies, the ways in which musicians make music in an ensemble is influenced by its participants and, if there is one, by the conductor's body. The role of body movements in conducting is self-evident when it comes to the conductor's hand gestures. Extensive literature has tackled the ways in which performers' movements influence the perception and the subjective experience of music [1–3], the functions of performers' movements, body communication among co-performers [4], body movements in singing [5,6] and in musical teaching [7]. Conversely, fewer studies have been devoted to the expressive use of face in performance, and research on the role of gaze in conducting is sparse. However, some musicians and musicologists cite the magnetic force of gaze of some conductors: for example, [8,9] the eyes of Antonio Guarnieri (1880–1952), who used very few and narrow gestures, but conducted through his sharp and penetrating gaze, that “bewitched” the orchestra.

Gaze is a very rich, complex and sophisticated communication system, and the parameters that determine its meaning are not only gaze direction or pupil dilation, but a number of physical aspects in the eye region, such as eyebrow movements, eyelid position, eye humidity and reddening, and the

like. This communication system can be studied in the same way as natural languages are, to the point that it is possible to write down a lexicon of gaze, and even an “optology” [10], equivalent to the “phonology” of verbal languages and to the “cherology” (the “phonology” of hand gestures) proposed [11] for the sign languages of the deaf. However, for the gaze communication system we use in everyday life, it is possible to write down a lexicon and an “optology”, therefore, why not write down a lexicon of the conductor’s gaze? Our hypotheses are: that a conductor in music performance uses gaze as one of the tools of his/her communicative job of conducting; that this gaze system, like the one laypeople use in everyday life, is a specific communication system, with systematic correspondences between eye actions and conveyed meanings; that, nonetheless, the specific gaze communication system used during conducting is partly similar and partly different from the one used in everyday life.

The aim of this work is to investigate if conductors are generally aware of the importance of gaze communication in orchestra and choir conducting, to demonstrate its relevance as a tool for effective conducting, to provide a sketch of a lexicon of the conductor’s gaze, and to test its systematicity and sharing among musicians and laypeople.

In Section 2, we overview previous works on gaze communication and the lexicon of gaze in everyday life, in musical performance in general, and then specifically in orchestra and choir conducting. Section 3 illustrates previous works on the lexicon of conductors’ gaze, and subsequent sections present three studies: an interview with five conductors to test their awareness of their own use of gaze as an instrument for conducting (Section 4), a qualitative analysis of Leonard Bernstein’s gaze communication during a concert (Section 5), and a perception study testing the comprehensibility of some of Bernstein’s communicative gaze items by music experts, amateurs, and non-experts (Section 6).

2. Gaze in Everyday Interaction

Ever since ancient rhetoric, Aristotle, Cicero, and Quintilian have acknowledged the relevance of gaze in the orator’s communication [12]. In modern psychological research, the seminal work of [13] overviewed the functions of gaze in man and other animals; Ekman [14] pointed out the expressive import of the eyebrows in conveying surprise and other emotions, but also anticipated their syntactic use as interrogative markers, which [15] later described as particularly precise and systematic in sign languages of the deaf. Further uses of gaze were illustrated, from conveying the performative of communicative acts [16,17], through regulating turn-taking and asking and giving backchannel, thus managing synchronization with speech and the negotiation of participant roles in interaction [18,19], fulfilling meta-discursive functions such as skipping a topic [20], and rhetorical functions like emphasizing or showing irony about something [21]. Nevertheless, eyes are not only exploited for expressive or discursive functions, as they can also bear referential information, for instance, pointing at objects or people in the surrounding context, but also producing iconic signals to inform about physical and metaphorical properties such as small or big size, subtle and difficult concepts [10]. However, gaze direction, which is so often taken into account in gaze studies, is not the only relevant parameter to imprint meaning into gaze communication. More generally, gaze can be seen as a lexicon: a set of systematic correspondences between meanings—i.e., various types of information on the world and on the Senders’ internal states—and signals—i.e., physical aspects of gaze. In writing down such a lexicon [10], distinguishing the above meanings of gaze (syntactic, referential, discursive and conversational functions) also proposes an analysis of gaze signals in terms of a set of physical parameters (gaze direction, eye position in the sclera, eye humidity and reddening, eyelid and eyebrow position and movements) with respect to the extent in which any gaze item may assume some possible values (e.g., eyes upward vs. downward in the sclera, half-open vs. wide open vs. closed eyelids, raised vs. frowning eyebrows). This allows us to account for the high flexibility of gaze in producing communicative signals and conveying an articulated lexicon of meanings in everyday communication, much richer than the one accounted for by only tracking people’s gaze direction.

2.1. Body and Gaze in Music Performance

The vast majority of works on the body communication of musicians or conductors investigates the use of gestures or whole-body movements. In the last 30 years, they have investigated the functions of movements in performance [22,23], sometimes focusing on the meanings of performers' movements [24–26], and how they influence the perception and the subjective experience of music [1–3,6,27], for instance, assessing the performer's gaze direction and other cues such as nodding, self-touch, stance width, step size, resolute impression in the first minute after stage entrance [28] and their effects on the audience's first impression. Other works focus on the reciprocal body communication between co-performers, such as progressive coordination, synchronization, and time lag reduction achieved through gestures and glances in duo pianists [4,29,30] and through head and gaze direction in quartets [31]. Concerning communication in ensembles, [32], through automated motion detection, shows how both head and gaze direction and head ancillary movements of a group of players spontaneously change when changing the way they interact with the rest of the orchestra and when they see the conductor. The study in [33], by mobile eye-tracking, finds a high variability in the amount of gazing at the partner in clarinet, violin and piano trios.

The study in [34] stresses the importance of gaze, along with sound, body movement, facial expression, and breath in giving and seeking information during musical interaction, and distinguishes mutual gaze—simply looking at the other's body—from eye-contact, which constitutes looking into each other's eyes. [35] finds that the players in a duo watch each other more after rehearsal than before, and that gazes are longer during temporally-unstable than during regularly-timed passages.

Overall, these results might indicate that a collaborative and monitoring gaze may be established only after a shared interpretation of the piece is determined. The study in [36] finds that the onset and duration of the looking behaviors of improvising instrumentalists does not co-vary with musical structure or the social context of performance; the underlying timing mechanisms support social interaction processes, facilitating interpersonal communication. In an experiment where subjects only saw or only heard other band members, [37] observed glances to the conductor 28% of the time, with glances under one second. The study in [38] also shows that pianists use nonverbal audio and visual cues during duet performances, while one in [39], which investigated several aspects of social interaction in piano duos, including dominance relationships, conflict management, and nonverbal communication, pianists highlight the importance of expressive body movements and positive facial gestures, but most of all of eye-contact which, gradually increasing from rehearsal to performance, is essential for coordination of the musical content.

The study in [40], besides from describing peculiar gestures, collects the musicians' descriptions of various types of eye contact among first and second violin and cellist, and argues that such "conversation with the eyes" also conveys the special social relation among performers.

On the other side of the stage, [31] found that for both non-musicians and musicians, gaze information helps to perceive musical expressivity. The study in [41] used eye-tracking to analyze the audience's gaze during a multipart female duo musical performance, finding that the melody part significantly attracted more visual attention than the accompaniment one; moreover, joint attention phenomena emerged as the singers shifted their gazes toward their co-performers, and melody or accompaniment strongly influenced the total duration of gazes within the audience. Comparing the utilization of acoustic cues of guitar pieces in conveying emotions, [42] found that performers were successful at communicating emotions to listeners, that their cue utilization was well matched to listeners, and that it was more consistent across different melodies than across performers.

2.2. The Conductor's Body

The social relationships in small ensembles and orchestra, and their mechanisms of leadership and group interaction have been the object of analyses ranging from phenomenological [43] to computational [44], all stressing the central role of the conductor, to the point where changing the conductor brings about a change in entertainment, and hence lower ratings of the performance [44].

Concerning the conductor's communication, gestures are the most frequently studied in the literature on conducting [22,45,46], while in the field of gesture studies, [47–49] analyze iconicity and metaphors in gestures. As to other modalities, [50] finds that eye contact and facial expressions of approval and disapproval determine a better opinion of more expressive conductors, irrespective of actual better performance. The study in [51], investigating the use of facial expressions in conjunction with musical conducting gestures and the relative interpretation by instrumental performers, found that only their experience level, but not the conductor's concomitant facial expression, affected the instrumental performers' ability to interpret 53 different musical conducting gestures. On the contrary, using an occlusion technique, [52] highlighted the importance of "facial affective behavior for expressive conducting". In his experiment, 127 participants ranging between musically trained and untrained, watched randomly presented video excerpts without sound, in which they could see only the face, only the arms or the whole body in simulated peripheral vision: the "only arms" condition was judged as the one conveying the most information, but the interpretation of excerpts showing the face was the most similar to one of the complete video sequences with sound taken as reference. Some studies ask conductors to self-assess their own potential for leadership or other conducting capacities [53], thus investigating their self-awareness of their own use of communicative cues and techniques.

When it comes to the specific use of gaze, more research is presently underway, also exploiting eye-tracking [54]; nonetheless, eye-gaze is mainly investigated in the audience, in performers and between co-performers [35], and rarely as a specific tool of the conductor's communication. In reality, some works address the relevance attributed by conductors to the various cues, including gaze. For instance, some consider eye contact so important that they recommend annotating the points in the score during which to gaze at musicians [55]. However, in all these cases, most typically, the parameter of gaze taken into account is almost exclusively gaze direction. On the contrary, what should be (but is not) generally acknowledged in the literature is that the range of communicative potentialities and corresponding physical patterns of gaze actions by conductors is much wider than bare glances. It is not simply the very fact of looking at performers that is meaningful; the conductor can resort to a whole repertoire of several different ways to use eyes, each consisting of a particular pattern of gaze actions and conveying a specific meaning to performers.

2.3. Body Lexicons of Musicians and Conductors

Some works concerning multimodal communication in musical performance evidence that the body behaviors of both musicians and conductors form proper lexicons, that is, sets of signals systematically corresponding to given meanings concerning specific aspects of music. In pianists, [26] found that the movements of the head, trunk, and face sometimes only have a motoric function, merely helping to perform the technical gesture, like when a circular movement of fingers on the piano is accompanied by a circular movement of the head. In other cases, body movements do not communicate anything to others but simply express the performer's internal state, e.g., *closed eyes* express concentration, *internal parts of eyebrows up* express sadness. As found in both pianists and conductors [26,56], four kinds of emotions are expressed during a performance: 1. Process Emotions, those felt while playing or conducting—tension, fear of making mistakes, excitation or flow; 2. Outcome Emotions, due to the appreciation of how the performance is going: shame or satisfaction, disgust or ecstasy; 3. Meaning Oriented Emotions, enacted by musicians or conductors in order to be impressed on the music being performed, e.g., sadness, joy, triumph, poignancy; 4. Movement Oriented Emotions, which expression implies (and helps to perform) a particular movement or manner of movement of the technical gesture. For example, both pianists and conductors typically *frown* when playing or asking for a "forte", because frowning is an expression of anger, and displaying—hence, somehow feeling—anger mobilizes the energy necessary to play loud. Finally, body movements may have an utterly communicative function in music performance since they convey specific meaning to others; this occasionally happens in players, but is always the case for conductors.

Therefore, various studies have attempted to find out the lexicons of the conductors' gestures and gaze: first, qualitative analyses of corpora of conducting hypothesize the correspondences between specific gestures or gaze items and the meanings conveyed by the conductor, and then perception studies test the hypotheses put forward about some gesture or gaze items. After singling out the gestures for intensity—those providing indications like *forte*, *piano*, *crescendo*, *diminuendo* [56–58]—a lexicon of gaze in conducting was proposed [59,60].

3. The Lexicon of the Conductor's Gaze

This work is embedded in a more general project aimed at finding out the repertoire of the conductor's body communication. The project sticks to a model that defines communication as any event in which a Sender (S) has the goal for an Addressee (A) to learn 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, through a Communicative System (CS), means a system of rules to set correspondences between signals and meanings, represented in both the Sender's and the Addressee's minds. The goal of communicating may be a conscious one, like when a conductor *closes his eyes* to signal that he is concentrating before starting—hence, that he is going to start; but it may also be a goal that S is not aware of, like when his *sweating* expresses his anxiety about the performance. The signals, according to the modality in which they are produced—words, prosody, gestures, facial expressions, gaze, posture—make up different Communication Systems, which may be of two different types. A “creative” system is a set of rules of resemblance used to create new iconic signals by imitating places, shapes, or actions: e.g., *curve arms widening*, imitating a swelling body, make an iconic gesture asking for “crescendo”. A “codified” system, that is, a “lexicon”, is a list of rules of correspondence in which a specific meaning corresponds to each signal in the Senders' and the Addressees' minds: e.g., *index finger as a vertical bar touching the mouth* is a gesture asking for “piano”; *frowning eyebrows* is a gaze item asking for “forte”.

Positing the existence of a lexicon of conducting gaze is based on the idea that if a signal used by a conductor did not correspond to the same meaning in both his/her mind and in the musicians', they simply could not understand each other, and the performance would be affected thereof. Our hypothesis is therefore that conductors and musicians share a lexicon of gaze in which each gaze item corresponds to a specific meaning; our previous works [59,60] and the present one first try to derive the items of this lexicon, and secondly try to test whether they are in fact shared across performers, and if they are also comprehensible to non-musicians.

In [60], in a corpus of conducting of an amateur choir, 17 items of gaze were singled out and described in terms of their parameters (e.g., gaze direction, eyelid and eyebrow positions, and movements) and literal and indirect meanings were attributed to each of them, finally clustering them into types according to the functions they fulfill in musical performance (Table 1). While some uses of gaze are non-communicative (e.g., line 3, the conductor *looks around* to check for the musicians' attention), some have an interactional function, such as requesting attention or providing feedback (line 7, he *opens his eyes wide* to reproach for a mistake), others have a technical function, providing indications for the various parameters of music: *raising eyebrows* to ask for a *piano* (line 11) or *squinting eyes* to ask for a “sforzato” (line 15) (indications of intensity), suddenly *gazing at a section* of the orchestra to give the start (line 1). Finally, an emotional/attitude category includes items that exploit gaze expressions of Meaning Oriented or Movement Oriented Emotions, either aimed at conveying expressive indications or to suggest the right attitude in playing or singing: e.g., *raising internal parts of eyebrows* to ask for a sad sound (line 13), *raising eyebrows while retracting head in the shoulders* to ask for accuracy (line 12) [59,60].

Table 1. The lexicon of the Conductor's gaze.

1 Gaze n.	2 Gaze Item	3 Literal Meaning	4 Possible Indirect Meaning	5 Function
1	<i>Gazes at X</i>	Request for attention	Prepare to start	Technical (start)
2	<i>Gazes around at all musicians</i>	Broadcast request for attention		Interactional
3	<i>Looks at all musicians</i>	Checking gaze. (Non-communicative)		Self-information
4	<i>Raised eyebrows with oblique gaze</i>	Warning gaze	I warn you about a difficult passage	Interactional
5	<i>Raised eyebrows with wide open eyes</i>	Emphasis	I ask for higher attention	Interactional
6	<i>Eyebrow frown with wide open eyes (+ extended index finger)</i>	Peremptory order		Interactional
7	<i>Wide open eyes fixing X</i>	Threatening gaze (to prevent similar behaviour)	I reproach you for your mistake	Interactional
8	<i>Raised eyebrows (+nodding)</i>	Appreciation + approval	I praise you	Interactional
9	<i>Continuous eyebrow frown (+ rocking head)</i>	Request to continue		Technical
10	<i>Short single eyebrow raising</i>	Higher note		Technical
11	<i>Raises eyebrows all along the musical fragment</i>	Imitation of light movement	Play/sing soft	Technical (intensity)
12	<i>Raises eyebrows (+ head in the shoulders)</i>	Caution gaze	Be accurate and precise	Attitude
13	<i>internal parts of eyebrows raised</i>	Sad gaze	Play/sing in a sad way	Emotional
14	<i>Frown</i>	Angry gaze	Feel/express anger → play aloud	Technical (intensity)
15	<i>Squints eyes</i>	Imitation of effortful movement	Play/sing "sforzato"	Technical (intensity)
16	<i>Closed eyes</i>	Concentration	I want (you) to enjoy the pleasure of music	Emotional (Motivational strategy: non-musical)
17	<i>Squeezed eyes (+ trunk retracting backward)</i>	Disgusted gaze	Outcome emotion → Neg. feedback	Interactional

Adapted from [59].

In the following, we present three studies, to try to deepen our knowledge about the use of communicative gaze in conducting. First, an interview submitted to five conductors to assess their awareness of the communicative importance of gaze during the performance; second, a qualitative observational study analyzing a concert by Leonard Bernstein conducted only by gaze and facial expression, aimed at singling out some of his gaze communicative items; third, a perception study testing the comprehension of some of these items by professional experts, amateurs, and non-experts.

4. Study 1. The Conductors' Awareness of Gaze Communication in Music Performance

As we have seen, some of the above studies [53,55] test the conductors' awareness of the cues, techniques, and strategies they exploit in conducting. In this work, before presenting, in Sections 5 and 6—two studies concerning the specific role of gaze in conducting—we want to test the conductors' level of awareness of their eye communication. What do conductors think about their

own use of gaze while conducting? Do they monitor their own use of eyes? And what importance do they attribute to it as a tool for communication in performance?

Here we present a study aimed at testing the degree of awareness, in conductors, of the importance of gaze communication during choir or orchestra conducting

4.1. Method

To investigate this issue, we submitted an in-depth interview to five professional conductors (All conductors asked to remain anonymous), in order to assess their use of gaze in conducting and their awareness of it. The first two interviews were aimed at establishing a control sample before focusing on some more specific questions in the subsequent three ones.

The nine questions of the interview had the aim of exploring the participants' previous thoughts about gaze communication, but also of testing their awareness of it in their conducting style and their inclination to give more value to such a neglected tool as gaze (Table 2).

Table 2. An interview on the use of communicative gaze by conductors.

(1) What aspects of music do you prefer to emphasize in your style of conducting? (intensity and dynamics, variations in time or rhythm, expressive elements, attention request or feedback etc.)
(2) Do you always prioritize the same elements regardless of the repertoire?
(3) Do you express your emotions during conducting? (What emotions? How do you express them? For what purpose?)
(4) During your conducting, what are the functions you use your eyes for?
(5) Among the aspects of music mentioned before (intensity and dynamics, variations in time or rhythm, expressive elements, request for attention or feedback, etc.), which ones, in your opinion, are better suited to be conveyed through gaze?
(6) From 1 to 10, how do you rate the importance of gaze while conducting?
(7) If you had to choose only one option, what main function would you attribute to your gaze during orchestra/choir conducting?
(8) What kind of input do you intend to convey through your gaze when you are conducting?
(9) How much of your intentions you communicate by your eyes do you think musicians/singers receive?

The first two questions of the interview were

(1) What aspects of music do you prefer to emphasize in your style of conducting? (intensity and dynamics, variations in time or rhythm, expressive elements, attention request or feedback, etc.)

(2) Do you always prioritize the same elements regardless of the repertoire?

aimed at understanding the conductor's style and to single out the elements he or she preferred to convey while conducting: for example, it would not be pertinent to extract data about rhythm if the conductor him/herself says s/he usually prefers to convey dynamics.

Question

(3) Do you express your emotions during conducting? (What emotions? How do you express them? For what purpose?)

generally investigated the relationship between the conductor and his/her own emotions.

The next two questions focused on the use of gaze in conducting.

(4) During your conducting, what are the functions you use your eyes for?

(5) Among the aspects of music mentioned before (intensity and dynamics, variations in time or rhythm, expressive elements, request for attention or feedback, etc.), which ones, in your opinion, are better suited to be conveyed through gaze?

(6) From 1 to 10, how do you rate the importance of gaze while conducting?

These questions tested the awareness of gaze in the five conductors, taking their experience, but also their preferences into account.

Question 4 highlights how consciously the conductor uses gaze to communicate with the performers, while Question 5 investigates his/her general perception of gaze as a tool that can be potentially used, even though s/he might not use it. Question 6 focuses instead on how much importance gaze assumes for the conductor, in relation to other communication channels, on a scale 1 to 10.

(7) If you had to choose only one option, what main function would you attribute to your gaze during orchestra/choir conducting?

(8) What kind of input do you intend to convey through your gaze when you are conducting?

Questions (7) and (8) try to assess if the conductors have conscious communicative intentions right behind their own use of gaze.

(9) How much of your intentions you communicate by your eyes do you think musicians/singers receive?

This last question is the only one in the list that is not about what conductors themselves think or feel about the gaze, but aims at understanding if conductors can judge the potential, and the result, of this communicative tool.

Every interview was conducted face to face. Although the interview was fairly structured, some questions had to be reiterated when the answer was not pertinent, mainly in order to overcome some language barriers. The interviews were audio-recorded and transcribed verbatim. The excerpts quoted below are translated from Italian.

4.2. Participants

The questions of the interview aimed at exploring the participants' previous thoughts. The choice fell on five professional conductors, namely, two from Italy, one from Lithuania, one from Argentina, and one from Bolivia; our perspective was as international as possible, since internationality is a crucial element to further highlight how gaze signals are culturally or linguistically determined. Second, to test whether awareness about gaze communication in conducting cuts across different conductors, we chose subjects of different gender and age (one female and four males, from 30 to 70 years old) and with different repertoires, two mainly conducting orchestra and three mainly choirs, ranging from classic to pop-modern music

4.3. Results

From the first questions of the in-depth interviews, two opposed trends emerge about conducting style: the tendency to emphasize the elements that convey emotions (preference for dynamics, phrasing, expressive elements) and, to the opposite, the tendency to emphasize the precision elements (starts and caesuras, time, rhythm).

"If I have to choose the moments in which I interact more with the orchestra, I would undoubtedly say when dynamics and phrasing are involved." (Conductor 1, Male)

The relationship between conducting style and repertoire is varied: three conductors claim that, while certainly taking into consideration the variables due to repertoire, all in all, the elements they tend to highlight are the same, while the other two highlight how repertoire is fundamental to set the conducting.

A significant element that deserves further study is that the conductors who claim to be tied to routine are those more connected to a choral dimension. A hypothesis to account for this lies in the type of work that the conductor carries out with a choir, which is often continuous (if not almost exclusive), or at least more assiduous than the one the conductor usually has with an orchestra (that is usually more tied to conventional signals).

"Precise starts, precise closures, well-timed, expressive sound. When you are working with a choir for a while, you also already know what weaknesses they could have and try to prevent them." (Conductor 2, Male)

As to the functions for which they use gaze signals, all the interviewed conductors seem to agree: answers mainly include the control and feedback functions and the notice of changes (both expressive and rhythmic). Interestingly enough, although they claim the main functions for gaze use for themselves are those of control, request for attention and feedback, in the aspects of music they, in general, believe gaze can effectively convey expressive elements and dynamics too—even if some conductors specify that they believe it to be unlikely that one can convey signals only through gaze, and they consider integrated communication more effective instead.

“Definitely the expressive elements and dynamics. Dynamics, however, together with the movements, because the eyes cannot give precise signs of when the phrasing should begin or end”.

They all agree, however, on the impossibility to transmit indications of time or rhythm only through gaze; they all believe the only thing that can be signaled is an imminent change (but the signal would equally lack the information necessary for performance).

“For time I prefer using hands, big movements. [...] there are pieces in which the choir sings without the conductor, in ancient music for example, but it is always one element of the choir who makes a sign to make others start, it is impossible for the choir to do it alone”. (Conductor 3, Male)

“Only through gaze? I don't really see how, you can't, you know, blink in time ... it would be pure madness”. (Conductor 4, Male)

Such ambiguity about the importance of gaze is also expressed by the rating (on a scale from 1 to 10) asked by Question 6. More specifically, the interviewed conductors rate it, respectively, 5.50, 6, 7, 7.75, and 9.50, thus ranging from skepticism to enthusiasm.

In summary, regardless of what intentions conductors claim they want to express through their eyes, all of them absolutely agree that gaze is definitely a clear signal for musicians because they are professionals and practice a lot before rehearsals.

“If only the gaze I don't know [how much they understand], together [with hands] surely they understand”. (Conductor 5, Female)

“As I said there was practice, every musician already knows what the conductor wants in that particular passage”. (Conductor 3, Male)

To sum up, the answers of the interviewed conductors concerning the importance of gaze and their actual use of gaze signals do not reveal thorough acknowledgment of the communicative import of this kind of signal in conducting. They claim that they make use of gaze mostly to ask for attention and give feedback and, as a secondary function, to convey information about emotions to performers; but by in large, they consider gaze signals mostly ineffective to convey indications related to time or rhythm, and in general, they believe that gaze is not autonomous as a conducting signal, but necessarily needs to concur with a gesture to convey meaningful indications.

Judging from these results, conductors on average do not seem to have a very high level of awareness of their own gaze communication, and of its relevance in conducting; or at least they do not seem to attach such concrete meanings to it and believe that gaze is more useful to trigger attention in general.

Nevertheless, although this kind of attitude towards gaze appears quite common in our interviews, we know, on the contrary, that some conductors are deeply aware of the importance, richness, and usefulness of this communication tool. Leonard Bernstein is one of them.

This is mainly witnessed by a performance of Haydn's Symphony No. 88 in G Major by Leonard Bernstein, with Wiener Philharmoniker in 1989 (The whole performance, lasting 29'51", can be viewed at [youtube.com/watch?v=AV_ZE4zcl3I](https://www.youtube.com/watch?v=AV_ZE4zcl3I). The fragment analyzed is from 25'15" to 28'52"), where during the encore after the performance he conducts with his hands down or behind his back and only using his face and gaze communication.

5. Study 2. Bernstein's Conducting Gaze—A Qualitative Analysis

Leonard Bernstein is one of the most complex and fascinating characters in the musical scene of the late twentieth century; a great orchestra conductor, an inspired composer, an effective pedagogue, and a controversial public figure. Despite his prolific authorial and executive production, what makes him the perfect subject for our research purposes is his tendency, especially in the last years of his career, to favor an intimate conducting style: at times his conduction has been made mainly by exploiting gaze and face, and the Haydn performance above is a transparent example of this. Bernstein is thus a perfect example to demonstrate that gaze in conducting can be credited the status of a lexicon.

5.1. Research Questions and Hypotheses

In this second study, we investigated Bernstein's use of gaze as a conducting tool, wondering if and what relevant indications for music performance are conveyed by Bernstein's eyes, and whether his items of gaze are understood by expert musicians and laypeople. To answer these questions, we carried out a qualitative analysis of his use of gaze during conducting.

Our hypothesis was that the types of gaze already found in previous analyses of other conductors [59,60] are also valid for Bernstein.

5.2. Materials and Method

To verify this, we carried out a qualitative analysis of a fragment of 3'37" from the above Haydn's Symphony No. 88 in G Major, conducted by Leonard Bernstein. This is a standard symphony composed of four movements, written for flutes, two oboes, two bassoons, two horns, two trumpets, timpani, basso continuo (harpsichord), and strings. We chose to analyze the *Finale (Allegro con spirito)* of this symphony for several reasons: first, Haydn is an author of the eighteenth-century, when the notation system was already sufficiently coded and the use of specific terms for dynamics and time was shared, so we have a solid basis for analyzing the conducting performance starting from the score; second, in the *Finale* all the aspects of music pointed at by the conductors in our interviews, such as dynamics, time, rhythm, are represented; third, thanks to the multiple tight shots and close-ups of Bernstein in the video, his gaze behavior is quite clear and can be analyzed better than with most orchestra performances.

The fragment was analyzed by two independent raters in terms of a simplified version of the annotation scheme proposed in [59,60] (Table 3). The analysis was carried out with muted video, to avoid being influenced by the corresponding sound, by describing and attributing a meaning to each gaze behavior considered as communicative; later, the ambiguous annotations were discussed between the two raters, and checked also with the corresponding audio.

Table 3. A fragment of the annotation scheme of Bernstein's gaze.

Time	Gaze	Meaning	Function
0.23	<i>Raises eyebrows</i>	Play Lighter	intensity
0.29	<i>Half-closed eyes</i>	I feel ecstasy, enjoyment → You are playing well	outcome emotion expression → Feedback
0.33	<i>Iris directed to left</i> + <i>Raised eyebrows</i>	I address you on my left alert → Prepare for start	attention start
0.38	<i>Closes eyes very fast</i>	Yes → ok	feedback
0.43	<i>Frowning eyebrows</i>	Play with determination	expressivity
0.47	<i>Half open eyes</i> + <i>Raised eyebrows</i>	I am serene → Play in a serene way	Meaning oriented emotion

In column 1, we write the time of the gaze item analyzed; column 2 contains the description of the item, 3 its meaning, and 4 its function for musical performance. In column 3, we write the literal meaning and sometimes, after the arrow, a possible indirect meaning which can be inferred from the literal one; finally, in column 4 we write the functions of the literal and, possibly, the indirect meaning written in 3, respectively. Thus, a single item may have two functions, one fulfilled by the literal meaning and one by the indirect meaning.

For example, at time 0.23 in our version of the video (column 1), Bernstein *raises his eyebrows* (column 2), a polysemous gaze item whose meaning common to all its uses [10] is “pay attention”; in this context this means “play lighter” (i.e., pay attention not to be too heavy in your movements) (column 3), with the function of conveying an indication of intensity (column 4).

5.3. Results

The annotation revealed that Bernstein’s gaze behaviors are similar to those performed by other conductors [59,60]. Like all of them, some signals have interactional functions (e.g., calling for attention, providing feedback) while others have technical functions (start, intensity, tempo, expression), and the emotion displays performed by gaze, again, like those found in previous analyses of other conductors, can express either the outcome of the ongoing performance (Outcome Emotion Expression), thus sometimes fulfilling a feedback function, or the emotion to be stamped in the music played (Meaning Oriented Emotion Expression). All in all, the fragment analyzed contains 56 gaze items, and the functions of their direct and indirect meanings are distributed as in Table 4.

Table 4. Distribution of gaze functions.

Gaze Function	Quantity
Attention	8
Feedback	10
Concentration	4
Start	16
Intensity	11
Pitch	1
Tempo	1
Duration	2
Expression	12
Outcome Em	5
Meaning Em	4
Total	74

As already undertaken in previous studies on intensity gestures [56–58], our method to discover the conductor’s lexicon exploits a two steps approach: first, we view the communicative signals investigated, whether gestures, or gaze items, only from the point of view of its production, hence simply attributing the Sender the intention of communicating a given meaning by using that particular signal. This was the aim of Study 2. The second step is to check, by a perception study, if that signal is actually interpreted as bearing that very same meaning by the Receivers; and this is what will be carried out in Study 3. In other words, in Study 2, we only make some hypotheses about what the meaning of each gaze signal in the fragment could be. Subsequently, the meanings hypothesized were tested in Study 3.

6. Study 3. How Comprehensible is Bernstein’s Lexicon of Gaze?

After deriving, through qualitative analysis, some of Bernstein’s conducting gaze items, we wanted to test their comprehensibility and sharedness in musicians and laypeople, and to this purpose we ran a perception study.

6.1. Research Questions

Our research questions were the following:

RQ1: How comprehensible are Bernstein’s gaze items?

RQ2: Is the level of comprehensibility in some way linked to the viewer’s musical expertise?

RQ3: Are they equally comprehensible if accompanied by corresponding sound or not? How is the interpretation of gaze samples affected by their being presented either in both audio and video modality or only visually?

RQ4: Do the most expert participants take greater advantage of the audio in recognizing the gaze items?

6.2. Experimental Design

To answer these questions, we designed a within-subjects perception study, with independent variables: 1. the meaning of gaze items, 2. the item presentation mode (video-only/audio-visual), and 3. the participant's musical expertise (Expert/Amateur/Non-expert), and as a dependent variable, the meaning attributed by the participant to the gaze item, considered as expected/unexpected with respect to the meaning resulting from our previous semantic analysis.

6.3. Materials

According to the opinions of the conductors interviewed in Study 1, the aspects of music performance that are most typically conveyed by gaze resulted to be the signals for "start", "keep the time", "pay attention", "crescendo/diminuendo", "accelerando/rallentando". Based on these results and on our analysis of Bernstein's performance, we chose three of his gaze items. For each of them, a video-clip was extracted from the analyzed fragment in order to show Bernstein's whole face, then the clips were presented either in audio-visual mode or in video-only mode without audio.

1. a gaze signal conveying the technical meaning "start": He looks at a section with open eyes, then he raises his eyebrows opening his eyes wide and then closes his eyes;
2. the "interactional" gaze signal for "pay attention": He opens his eyes wide while raising his eyebrows, and turns his gaze from right to left by moving his irises right-left in the sclera;
3. a gaze signal meaning "I ask you to perform a crescendo and accelerando": He raises his eyebrows while corrugating his forehead and then frowns intensely looking at the orchestra.

This last gaze item is quite a complex example, and it was chosen to demonstrate that even difficult items such as this are generally attributed meaningful interpretations.

As specified above, the clips showed the whole face of the conductor. An alternative would have been to cut the stimuli only showing his eye region. As demonstrated by [52] using occlusion techniques, the face is generally the richest part of the conductor's body in conveying expressiveness, i.e., information concerning the emotions to be stamped onto music. However, the meanings we attributed to the chosen gaze stimuli were not "emotional" but technical or interactional ones. In reality, in the stimuli "start" and "pay attention", the emotional import contributed by facial muscles outside the eye region did not bear relevant differences in the meaning of the gaze stimuli per se. In the "accelerando" part of the third stimulus, it is somehow the other way around: an originally emotional gaze signal—here a *deep frown*, generally conveying anger or concentration—comes to bear a non-emotional meaning—mobilize your energy to hurry up.

6.4. Pilot Study

We first ran a pilot study aimed at refining the procedure for the subsequent perception study. A focus group was conducted recruiting 12 participants, four for each of three levels of musical competence: Experts, Amateurs, and Non-experts. We considered people professionally working with music as "Experts": musicians, singers, composers and arrangers, conductors, music teachers; "Amateurs", those with some knowledge of music technique but not accustomed to the ritual of professionals: music students, non-professional musicians; and "Non-experts", those with no technical knowledge of music.

In the focus group, the clips of the three selected gaze signals cut from Bernstein's video were shown to the 12 participants in video-only mode while asking, for each of them, to guess its possible

meaning. The meanings proposed for each of the three items were recorded, to later construct the distractors for the subsequent perception study.

6.5. Method

A questionnaire was finally built up to test the participants' attribution of meanings to the three clips of Bernstein's video—the gaze items respectively meaning, according to our hypothesis, “start”, “pay attention”, and “crescendo and accelerando”. After answering biographical questions on gender and age, each participant watched the three clips of conducting behavior, each repeated twice, the first time in video-only mode and the second in audio-visual mode; then s/he had to answer a multiple-choice question for each clip, where the expected meaning was mixed with four distractors emerged from the focus group, not only the most quoted alternatives but sometimes the most interesting ones. All alternatives were randomized; in the video-only and the audio-visual version of the same clip, the answers were the same but in different random order. Finally, in the last question, following the audio-visual clip “crescendo/accelerando”, the participants were asked to rate, on a 10 step Likert-type scale, how certain the participant felt about the meaning of the previous clip. This question was put to participants regarding the third clip only, due to its particular complexity.

Before entering the questionnaire, the recruited participants were asked to classify themselves as Experts, Amateurs, or Non-experts, with the three classes defined as seen above. Finally, a control question asked them what their work or activity was, and those defining themselves Experts who did not do a musical job were considered Amateurs.

6.6. Participants

The sample, composed of 186 subjects, is quite balanced for age, with a slight drop in the “over 65”: 28% of participants between 18 and 25 years old, 22% between 26 and 35; 26% between 36 and 50; 18% between 51 and 64 years old; 6% over 65. As for gender, 54% participants are women. As for level of expertise, 37% are “Experts”, out of which 46% are musicians, 23% singers, 15% composers and arrangers, 9% conductors, and 6% music teachers; 29% are “Amateurs”, 34% “Non-experts”.

6.7. Results

For all statistical analyses, IBM SPSS 26.0 was used, and the significance level was set to $\alpha = .05$. Coherently, also observed power ($1-\beta$) was computed using $\alpha = 0.05$.

RQ1: Our first research question concerned the comprehensibility of the three gaze items by Bernstein submitted to participants. Our analysis indicates that they are fairly comprehensible. The overall percentage of correct interpretations without audio is 48% (For the sake of clarity, since we describe both the overall results and those of the single signals (§ 6.7.1-3), we sometimes employ the percentage format, some other times the average score of recognized signals; of course, they are interchangeable: if a single signal is recognized 50% of the times, this means that the correspondent average will be 0.5; as for the overall scores, 50% corresponds to 1.5 (out of three signals)). Considering that each answer concerning an item was chosen among five possible ones (one expected answer and four distractors), this is more than twice the level of chance.

RQ2: The second question was whether the level of comprehension depends on the viewer's musical expertise. Our results show a strict relationship between the viewers' musical expertise and the correct interpretation of the gaze items. A three-way analysis of variance (ANOVA) was performed with expertise, gaze signals, and modality of presentation as the factors and the correct interpretations as the dependent variable. The main effect of the expertise resulted to be significant, $F(2, 1104) = 28.34, p < 0.001, \eta_p^2 = 0.049, (1-\beta) > 0.99$. Subsequent post-hoc comparisons with Bonferroni correction showed significant differences between Experts ($M = 0.65$) and Non-experts (0.43) ($p < .001$), and Experts and Amateurs ($M = 0.43$) ($p < 0.001$).

As shown by Figure 1, Experts in both conditions (video-only and audio-visual) systematically perform better in the interpretations of the gaze items, followed by Amateurs (albeit with a

remarkable difference between interpretation from video-only and from audio-visual) and Non-experts.

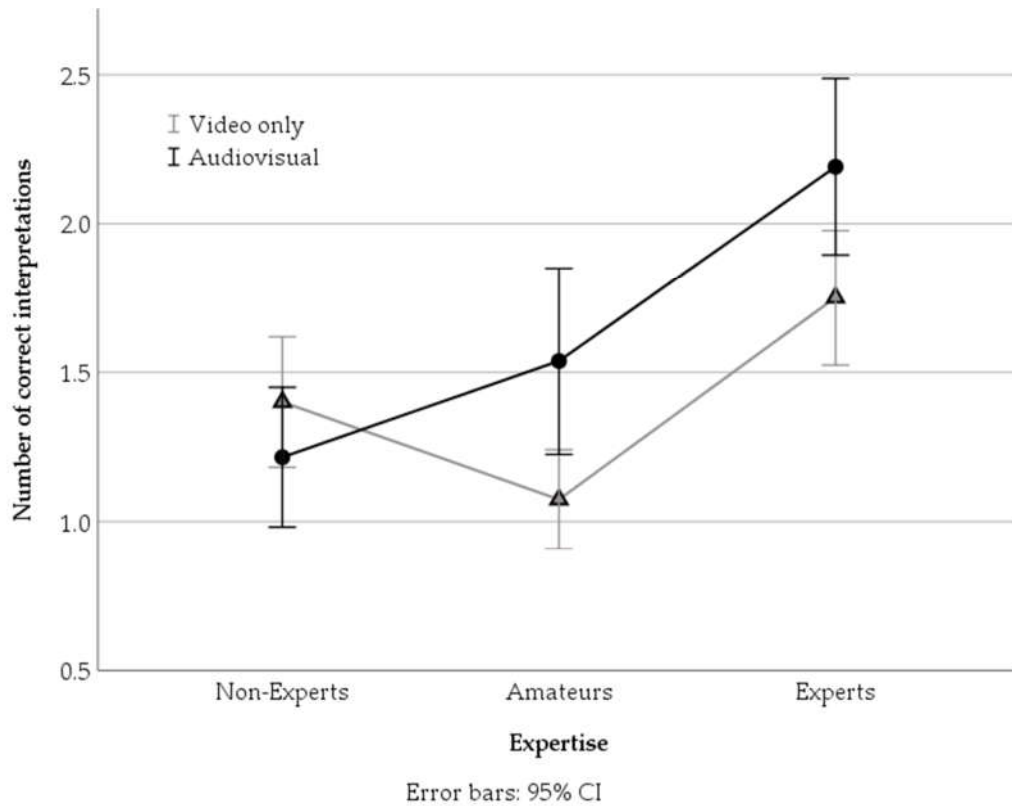


Figure 1. Correct interpretations as a function of expertise.

RQ3: The third question was about the possible differences in item comprehension depending on the stimuli being accompanied by audio or not. We found a main effect of the modality, namely, not surprisingly, the audio-visual condition ($M = 0.55$) led to a better understanding of the gaze items as opposed to the video-only condition ($M = 0.47$), $F(1, 1104) = 7.91$, $p = 0.005$ $\eta_p^2 = 0.007$, $(1-\beta) = 0.80$.

RQ4: The three-way ANOVA revealed that there is a significant interaction between Expertise and the modality of presentation, $F(2, 1104) = 6.36$, $p = 0.002$ $\eta_p^2 = 0.01$, $(1-\beta) = 0.90$: in other words, the differences between video-only and audio-visual presentation significantly vary among our three experimental sub-samples. Bonferroni corrected post-hoc analyses revealed that Amateurs and Experts, as opposed to Non-experts, took advantage of the audio for correctly guessing the right meaning. Post-hoc results are reported in Table 5.

Table 5. Pairwise comparisons in gaze comprehensibility across expertise and presentation modality.

Expertise	(I) Modality	(J) Modality	Mean Difference (I-J)	Std. Error	<i>p</i> .
Non-Experts	Video only	Audiovisual	0.062	0.048	0.200
	Audiovisual	Video only	-0.062	0.048	0.200
Amateurs	Video only	Audiovisual	-0.154 *	0.053	0.003
	Audiovisual	Video only	0.154 *	0.053	0.003
Experts	Video only	Audiovisual	-0.147 *	0.047	0.002
	Audiovisual	Video only	0.147 *	0.047	0.002

p-values refer to Bonferroni correction of the three-way ANOVA post-hoc analysis. * The mean difference is significant at the 0.005 level.

Moreover, if we take into account the different gaze signals, we discover that Amateurs performed significantly better with the addition of the audio in “pay attention” ($p = 0.005$) and “crescendo and accelerando” ($p = 0.008$); while Experts performed better only in “crescendo and accelerando” ($p < 0.001$) as highlighted by the significant interaction Expertise \times Modality \times Gaze, $F(4, 1104) = 3.46$, $p = 0.008$ $\eta_p^2 = 0.01$, $(1-\beta) = 0.86$.

We also took a closer look at the single gaze signals to search for any differences among them. In general, as hypothesized, the “crescendo and accelerando” was the less recognized one ($M = 0.41$) as opposed to “pay attention” ($M = 0.53$) and “start” ($M = 0.58$), $F(2, 1104) = 13.21$, $p < 0.001$ $\eta_p^2 = .02$, $(1-\beta) = 0.99$. Subsequent post-hoc comparisons with Bonferroni correction confirmed this result, in that “crescendo and accelerando” significantly differed both from “start” ($p < 0.001$) and “pay attention” ($p = 0.001$).

The difference in the correct interpretation of the stimuli is also proved by the significant interactions between Modality \times Gaze signals and Expertise \times Gaze signals, $F(2, 1104) = 8.24$, $p < 0.001$ $\eta_p^2 = 0.01$, $(1-\beta) = 0.96$, and $F(4, 1104) = 2.84$, $p = .023$ $\eta_p^2 = 0.01$, $(1-\beta) = 0.77$.

As can be observed in Figure 2, when the audio is added, all the three groups behave similarly; on the contrary, in the video only condition, the Amateurs performed poorly in recognizing the “pay attention” and “crescendo and accelerando” signals. In both modalities, the scores of the Non-experts, as opposed to the others, remain stable.

Curiously enough, by means of a one-way ANOVA, we also found a gender effect for the Video-only condition, $F(1, 185) = 5.46$, $p < 0.021$ $\eta_p^2 = 0.03$, $(1-\beta) = 0.64$, namely men ($M = 1.59$ $SD = 0.80$) tended to recognize the signals better than women ($M = 1.30$ $SD = 0.91$). To be on safer side, we also checked for the gender distribution among the levels of expertise, finding it to be not significant, $\chi^2(2, N = 186) = 3.47$, $p = 0.17$, thus concluding that, at least as far as gaze conducting signals are concerned, males perform better than females. These data deserve further and more focused investigation since such higher proficiency in body language of men appears to contradict not only popular wisdom, but also several findings of a female advantage in social perception tasks [61,62], among which facial emotion recognition test [63] and recognition of distinct features depicting bodily movements [64]. However, this difference might be confined to technical conducting signals only.

We now proceed to a more fine-grained description of each single gaze signal.

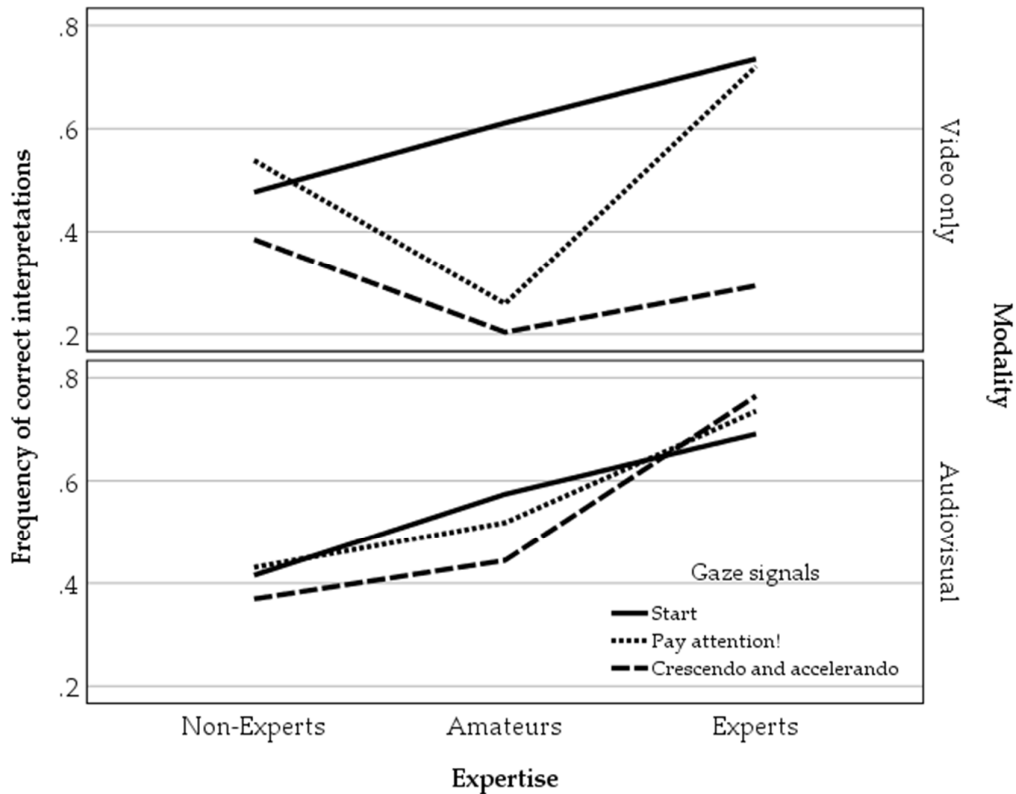


Figure 2. Correct interpretations as a function of expertise, modality, and gaze signals.

6.7.1. Gaze Signal for “Start”

The first video represented the gaze signal that indicates “start”.

For its recognition the gap between right interpretations from audio-visual and video-only presentation is wider as the level of expertise decreases. While Experts recognize the signal even without audio ($M = 0.073$ $SD = 0.44$), this mean drops, in a fairly predictable manner to $.61$ ($SD = 0.49$) for Amateurs and to $.47$ ($SD = 0.50$) for Non-experts, the only significant mean difference being that of Experts and Non-experts ($p = 0.005$); the pattern remains the same in the audio-visual mode, with Experts being at $.69$ ($SD = 0.46$) and Non-experts at $.41$ ($SD = 0.49$) ($p = 0.002$).

Although differences in recognition of the signal derive from the different music expertise, all in all the signal is understood on average by 61% of the sample, in the case of video-only clips, and by 56% when sound is added.

6.7.2. Gaze Signal for “Pay Attention”

The second clip is the one with the gaze-signal for “pay attention”. Its meaning is peculiar since it expresses the relationship that exists between the conductor and the orchestra. This peculiarity seems to be confirmed by our data, in that the majority of Experts ($M = 0.72$ $SD = 0.45$) recognize the “pay attention” signal even in the absence of audio; this is an indication that, at least for a competent audience, the signal is highly codified and shared. On the contrary, Amateurs recognize it less, although just above the chance level ($M = 0.24$ $SD = 0.44$), whereas Non-experts correctly interpret it more frequently ($M = 0.53$ $SD = 0.50$). Bonferroni corrected post-hoc comparisons reveal that Amateurs significantly differ both from Experts ($p < 0.001$) and Non-experts ($p = 0.004$). When audio is added, while Experts’ correctness does not increase ($M = 0.74$ $SD = 0.44$), Amateurs improve their performance up to $.52$ ($SD = 0.50$) ($p = 0.005$) and Non-experts perform worse ($M = 0.43$ $SD = 0.49$), although not significantly.

In reality, the most interesting result about this gaze signal comes from the comparison between Amateurs' and Non-experts' responses. In both cases, introducing the auditory stimulus modifies participants' recognition: with Amateurs, the auditory stimulus helps to discern reality, to identify the correct meaning (they double their rate of correct responses), while the case of Non-experts is quite different.

The data obtained from Non-experts' answers, at first sight, seem contradictory. Unexpectedly, as much as 53% of Non-experts (almost ten percentage points more than the interpretation of the gaze-signal for "start", which is, in general, the most recognizable signal) identify the right option even without audio, but this percentage drops (43%) when the auditory stimulus is introduced in the second version of the video.

Overall, this signal is understood on average by 52% of the sample, in the case of video-only clips, and by 56% when sound is added.

6.7.3. Gaze Signal for "Crescendo and Accelerando"

The third video clip explored the comprehensibility of a gaze signal issued by Bernstein that was in fact quite complex: one asking for "crescendo and accelerando". This indication is complex because it mixes two indications pertaining to different parameters of music: intensity (crescendo) and tempo (accelerando). Among our participants, Experts are able to recognize the indication of intensity; a positive response by .29 (SD = 0.45) (they are not able to perceive the indication of tempo combined with that of intensity). Amateurs even less than that (M = 0.20 SD = 0.40). Surprisingly enough, Non-experts managed to recognize it almost twice as much as the Amateurs (M = 0.38 SD = 0.49). None of these differences are statistically significant as shown by Bonferroni corrected post-hoc comparisons.

Things radically change with the addition of sound. The "perform crescendo and accelerando" option for Experts significantly rises up (M = 0.76 SD = 0.42) ($p < 0.001$), revealing that this population takes advantage of the audio more than the other sub-samples. Amateurs take a great deal of advantage too (M = 0.44 SD = 0.50) ($p = 0.008$); Non-experts' ratings remain low (M = 0.37 SD = 0.48), but always above the level of chance.

On average, this signal is understood by 30% of the sample, in the case of video-only clips, and by 53% when sound is added.

6.7.4. Confidence Ratings

Given the complexity of this last gaze indication, "crescendo and accelerando", we decided to add a certainty self-assessment scale to rate the level of confidence of participants in their response: the participants were asked to identify, on a Likert scale from 1 to 10, the degree of confidence of their answer on the video administered. An omnibus one-way ANOVA was performed, revealing that the three groups significantly differ, $F(2, 184) = 33.82$, $p < 0.001$ $\eta_p^2 = 0.27$, $(1-\beta) = 0.99$. Bonferroni post-hoc analyses show that Experts (M = 7.50 SD = 1.32) significantly differ both from Amateurs (M = 5.05 SD = 2.26, $p < 0.001$, SE = 0.33) and Non-Experts (M = 5.29 SD = 1.96, $p < 0.001$ SE = 0.32) (Figure 3).

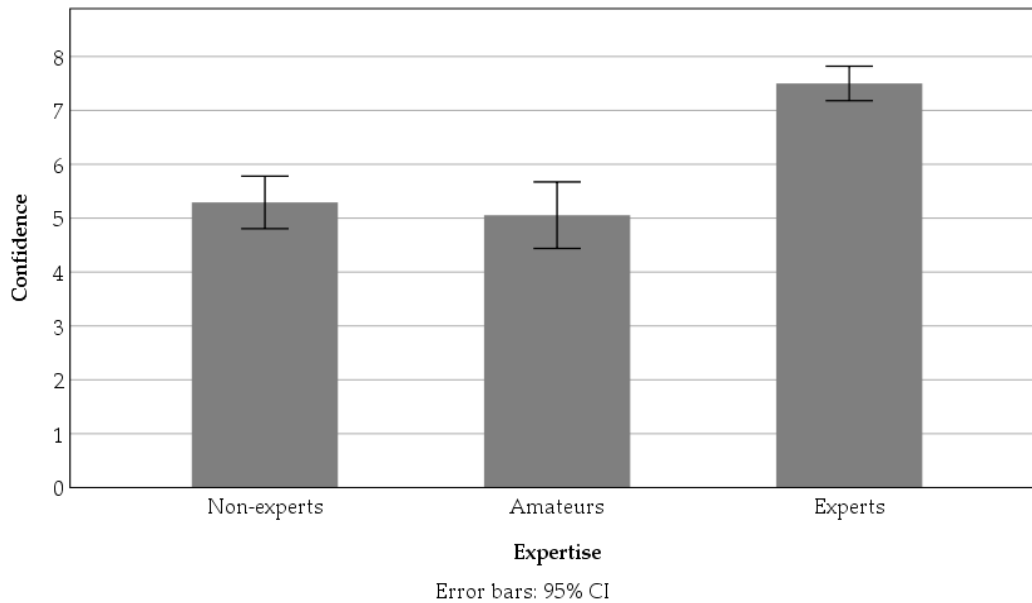


Figure 3. Confidence as a function of expertise.

6.8. Discussion

The hypothesis of this work was that the conductor's gaze could be encoded in specific signals, intelligible to some extent; our specific aim was to understand how recognizable are the signals of gaze used by Bernstein in his conducting, if their recognition is related to their level of music expertise, and whether it changes depending on the presence or absence of sound.

As for the first research issue, only the “start” gaze signal is fairly recognized by the whole sample. On the contrary, the signal for “pay attention” is less recognized. The delicacy of such signal is that by means of it the conductor transmits his emotions to the orchestra, he sends signals so that the performance reflects his idea of the work, and at the same time asks the orchestra for a commitment to attend to his signals and interpret them correctly; quite a complex relationship to understand for a Non-expert. A possible—yet curious—account for this result might thus be that signals like requests for attention and feedback (praise and reproach), that belong to the class of interactional signals, are not specific for conducting, so Non-experts do not expect to find them in a musical performance, and rather try to interpret them as technical signals in the presence of music; Experts and Amateurs, instead, having some experience of the relationship between conductor and performers, find these meanings rather expectable and recognize them both in the presence and in the absence of audio. In this perspective, it is not surprising that the “pay attention” gaze signal is difficult to understand for an audience totally uninformed about musical techniques: the sound is no longer an auxiliary element to understand, as was the case for amateurs, but it is an obstacle, a distracting factor.

With regard to the last and most complex signal, “crescendo and accelerando”, the interesting finding regards the participants’ interpretation confidence, and the slightly lower confidence of Amateurs as opposed to Non-experts.

Coming to the second research question, a significant difference emerges in the correctness of “video-only” interpretations among all three groups of participants. Experts systematically performed better than both Amateurs and Non-experts, whereas Amateurs, contrary to the hypothesis, provide less correct interpretations than Non-experts. When audio is added, instead, this strange effect disappears, and recognition linearly correlates with expertise.

Regarding our third and last research question, a difference emerges in correct interpretations of the video-only vs. audio-visual interpretations, namely the same signals with audio are better interpreted than those without it. Furthermore, the differences between with and without audio

significantly vary depending on expertise; in other words, Non-experts are the only participants who do not take advantage of the audio cue. This counts as an implicit confirmation of the importance of expertise to grasp the right meaning of the gaze signals.

This study on Bernstein's conducting gaze items can be compared with a previous work on the recognition of gaze items by the conductor of an amateur choir [59]; a common result is the high level of recognition of the item for "start"; but apart from this single item, in general, no significant difference had been found in that work between Experts' and Non-experts' recognition, whereas in the present study, Experts are better than both Non-experts and Amateurs. This might depend on the different gaze items (apart from the "start" one) investigated in the two works: in the previous one, the gaze items mainly concerned technical aspects of music, like tempo and intensity, whereas here an "interactional" signal was submitted to participants, which, as seen above, might have more clearly set the difference between Experts, Non-experts and, most of all, Amateurs.

7. Conclusion

During the last few decades, a conspicuous body of research has investigated multimodal communication in music performance, tackling the body movements of performers and conductors between each other and with the audience. Concerning gestures in conducting, they have been investigated including their semantic aspects, the meanings they convey and the semiotic devices they exploit [45,47–49,56–58]. In regards to gaze, the majority of studies on the use of eyes in performance, both between co-performers and between performers and conductor, deals with what we would call an "input", not an "output" function: using eyes mainly to acquire information about what another is doing [10], in order to reach synchronization or to catch technical or expressive indications. Even in studies tackling the "output", however, definitively communicative functions of the conductor's eyes do not catch the true richness of the conductor's gaze in conveying indications to musicians; even sophisticated studies through eye-tracking techniques, in music performance [33] like in everyday interaction [19], do not tell us much about the semantic input of gaze. First, the results are often confined to a single parameter such as gaze direction, without taking into account many other parameters such as eyebrow and eyelid movements or iris position on the sclera, that make up a rich repertoire of gaze signals [10,60]. Second, they do not even try to find systematic correspondences between the various signals and specific meanings for conducting.

The goal of our work, instead, was to demonstrate that there is much more than gaze direction to the communicative use of eyes in conducting: after discovering a shared lexicon in other conductors [59,60], we tried to find instances of this lexicon in the gaze communication of Leonard Bernstein. To this end, we conducted three studies exploiting three different empirical methods. First, an interview with five conductors told us that they are, to some extent, aware of their own use of gaze in conducting, but they do not attribute specific meanings to particular items of gaze, and they do not think that some parameters of music, such as rhythm, can be effectively communicated by gaze only. Second, an observational study coded the communicative gaze items used by Leonard Bernstein while conducting Haydn's symphony No. 88. Third, a perception study tested the comprehensibility of three of these items by Experts, Amateurs, and Non-experts when presented by a video-only vs. an audio-visual clip, finding out that the three items are reasonably recognized, but recognition is affected by both expertise level and presentation mode.

This work seems to confirm that the communicative items of gaze used in conducting make up a shared, systematic, and specific lexicon. The fact that, generally, Experts, and sometimes Amateurs, are better than Non-Experts in recognizing the items meanings tell us they in a sense make part of a "conducting gaze" communication system.

If some conductors, like Antonio Guarnieri quoted above, or Richard Strauss cited by [52], did in fact "conduct with their eyes", Leonard Bernstein did so in a conscious way: the very fact that, after conducting the whole symphony with the baton, in the encore, he decided to keep his hands down or behind his back, and to conduct by face and gaze only means that he was utterly aware of their self-consistency and even their effectiveness as a conducting tool.

Coming to the limitations of this study, we must keep in mind that gaze is only one of the signals produced by the face, while other parts of it—e.g., the mouth—may at the same time convey different messages. Therefore, we were very selective for choosing fragments to be used as stimuli in ways that other parts of the face did not add relevant information, nor intensify the meaning conveyed by gaze. Even then, the extent that the whole facial coordination contributes to the meaning attributed to a gaze item is not clear. Therefore, future research should provide refinements on this issue by submitting to participants clips with only the eye region visible in the video frames. Other studies will finally test the comprehensibility of a broader set of gaze items.

What is the point of such work? Studying the forms and meanings of the conductor's lexicon of gaze and the production and comprehension of its items may contribute, on the theoretical side, to in-depth knowledge of the communicative devices implied in music performance, and of their similarities and differences from those of everyday communication; on the application side, this may enhance the conductors' awareness of their communicative instruments and the teaching of conducting; on the technological side, to build virtual conducting devices, examples of which are conductor embodied agents and robots [65].

All of this work, finally, might be exploited to compare the effects on ensemble performance depending on the extent in which conductors make use of their gaze lexicon, and thus to understand in detail how the conductor's body behavior influences music.

Author Contributions: Conceptualization, I.P. and L.R.; Data collection and analysis, L.R., Y.L. and A.A.; Formal analysis, A.A.; Funding acquisition, I.P.; Methodology, I.P. and A.A.; Supervision, I.P.; Writing—original draft, I.P.; Writing—review & editing, I.P., L.R., Y.L. and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the Italian MIUR, PRIN “Cultural Heritage Resources Orienting Multimodal Experiences (CHROME)”, grant number 2015WXBPYK, and partially by a scholarship of the Department of Philosophy, Communication, and Performing Arts, n. 1236 – 04.09.2019.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Davidson, J.W. Visual perception of performance manner in the movements of solo musicians. *Psychol. Music* **1993**, *21*, 103–113.
2. Davidson, J.W. Which areas of a pianist's body convey information about expressive intention to an audience? *J. Hum. Mov. Stud.* **1994**, *26*, 279–301.
3. Davidson, J.W. Bodily movement and facial actions in expressive musical performance by solo and duo instrumentalists: Two distinctive case studies. *Psychol. Music* **2012**, *40*, 595–633.
4. Williamon, A.; Davidson, J.W. Exploring co-performer communication. *Musicae Sci.* **2002**, *6*, 53–72.
5. Davidson, J.W. The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. *Musicae Sci.* **2001**, *5*, 235–256.
6. Davidson, J.W. Qualitative insights into the use of expressive body movement in solo piano performance: A case study approach. *Psychol. Music* **2007**, *35*, 381–401.
7. Simones, L.L.; Rodger, M.; Schroeder, F. Communicating musical knowledge through gesture: Piano teachers' gestural behaviours across different levels of student proficiency. *Psychol. Music* **2015**, *43*, 723–735.
8. Piamonte, G. Ricordo di Antonio Guarnieri. *Musica* **1981**, *20*, 23.
9. Mandelli, A. *Antonio Guarnieri*; Edizioni MC Musica Classica: Milano, Italy, 1997.
10. Poggi, I. *Mind, hands, Face and Body: A Goal and Belief View of Multimodal Communication*; Weidler: Berlin, Germany, 2007.
11. Stokoe, W.C. *Sign Language Structure*; Linstok Press: Silver Spring, MD, USA, 1978.
12. Pallicer, M.A.F.; Rodríguez-Escalona, M.P. Mirar de reajo y fijar la mirada en los textos latinos*/Looking sideways and staring in Latin texts. *Cuadernos de Filología Clásica. Estudios Latinos* **2011**, *31*, 213.
13. Argyle, M.; Cook, M. *Gaze and Mutual Gaze*; Cambridge University Press: Cambridge, UK, 1976.
14. Ekman, P. About brows: Emotional and conversational signals. In *Human Ethology*; von Cranach, M., Foppa, K., Lepenies, W., Ploog, D., Eds.; Cambridge University Press: Cambridge, UK, 1979; pp. 169–248.

15. Thompson, R.; Emmorey, K.; Kluender, R. The relationship between eye gaze and verb agreement in American Sign Language: An eye-tracking study. *Nat. Lang. Linguist. Theory* **2006**, *24*, 571–604.
16. Poggi, I.; Pelachaud, C. Performative faces. *Speech Commun.* **1998**, *26*, 5–21.
17. Poggi, I.; Pelachaud, C.; De Rosis, F. Eye communication in a conversational 3D synthetic agent. *AI Commun.* **2000**, *13*, 169–181.
18. Jokinen, K.; Furukawa, H.; Nishida, M.; Yamamoto, S. Gaze and turn-taking behavior in casual conversational interactions. *ACM Trans. Interact. Intell. Syst.* **2013**, *3*, 1–30.
19. Brône, G.; Oben, B.; Jehoul, A.; Vranjes, J.; Feyaerts, K. Eye gaze and viewpoint in multimodal interaction management. *Cogn. Linguist.* **2017**, *28*, 449–483.
20. Vincze, L.; Poggi, I. Communicative functions of eye closing behaviours. In *Analysis of Verbal and Nonverbal Communication and Enactment. The Processing Issues*; Springer: Berlin, Germany, 2011; pp. 393–405.
21. Poggi, I.; Vincze, L. Gesture, gaze and persuasive strategies in political discourse. *The International LREC Workshop on Multimodal Corpora*; Springer: Berlin, Germany, 2008; pp. 73–92.
22. Godøy, R.I.; Leman, M. *Musical Gestures: Sound, Movement, and Meaning*; Routledge: London, UK, 2010.
23. Jensenius, A.R. Methods for studying music-related body motion. In *Springer Handbook of Systematic Musicology*; Bader, R., Ed.; Springer Handbooks; Springer: Berlin/Heidelberg, Germany, 2018; pp. 805–818.
24. Wanderley, M.M.; Vines, B. Origins and functions of clarinetist's Ancillary Gestures. *Music Gesture* **2006**, *167*, 165–191.
25. Clarke, E.F.; Davidson, J.W. The body in music as mediator between knowledge and action. In *Composition, Performance, Reception: Studies in the Creative Process in Music*; Thomas, W., Ed.; Ashgate Publishing: Aldershot, UK, 1998; pp. 74–92.
26. Poggi, I. Body and mind in the pianist's performance. In *Proceedings of the 9th International Conference on Music Perception and Cognition*; Alma Mater Studiorum University of Bologna: Bologna, Italy, 1 January 2006; pp. 1044–1051.
27. Vines, B.; Krumhansl, C.; Wanderley, M.; Levitin, D. Cross-modal interactions in the perception of musical performance. *Cognition* **2006**, *101*, 80–113.
28. Platz, F.; Kopiez, R. When the first impression counts: Music performers, audience and the evaluation of stage entrance behaviour. *Musicae Sci.* **2013**, *17*, 167–197.
29. King, E.; Ginsborg, J. Gestures and glances: Interactions in ensemble rehearsal. In *New Perspectives on Music and Gesture*; King, E., Ed.; Routledge: London, UK, 2016; pp. 203–228.
30. Kawase, S. Gazing behavior and coordination during piano duo performance. *Atten. Percept. Psychophys.* **2014**, *76*, 527–540.
31. Glowinski, D.; Riolfo, A.; Shirole, K.; Torres-Eliard, K.; Chiorri, C.; Grandjean, D. Is he playing solo or within an ensemble? How the context, visual information, and expertise may impact upon the perception of musical expressivity. *Perception* **2014**, *43*, 825–828.
32. Gnecco, G.; Badino, L.; Camurri, A.; D'Ausilio, A.; Fadiga, L.; Glowinski, D.; Sanguineti, M.; Varni, G.; Volpe, G. Towards automated analysis of joint music performance in the orchestra. In *Arts and Technology*; De Michelis, G., Tisato, F., Bene, A., Bernini, D., Eds.; Springer: Berlin, Germany, 2013; Volume 116, pp. 120–127.
33. Vandemoortele, S.; Feyaerts, K.; Reybrouck, M.; De Bièvre, G.; Brône, G.; De Baets, T. Gazing at the partner in musical trios: A mobile eye-tracking study. *J. Eye Mov. Res.* **2018**, *11*, 1–13.
34. Kawase, S. Importance of communication cues in music performance according to performers and audience. *Int. J. Psychol. Stud.* **2014**, *6*, 49–64.
35. Bishop, L.; Cancino-Chacón, C.; Goebel, W. Eye gaze as a means of giving and seeking information during musical interaction. *Conscious. Cogn.* **2019**, *68*, 73–96.
36. Moran, N. Improvising musicians' looking behaviours: Duration constants in the attention patterns of duo performers. In *Proceedings of the 11th International Conference on Music Perception and Cognition (ICMPC11)*, Seattle, WA, USA, 23–27 August 2010; pp. 565–568.
37. Fredrickson, W.E. Band musicians' performance and eye contact as influenced by loss of a visual and/or aural stimulus. *J. Res. Music Educ.* **1994**, *42*, 306–317.
38. Bishop, L.; Goebel, W. When they listen and when they watch: Pianists' use of nonverbal audio and visual cues during duet performance. *Musicae Sci.* **2015**, *19*, 84–110.
39. Blank, M.; Davidson, J.W. An exploration of the effects of musical and social factors in piano duo collaborations. *Psychol. Music* **2007**, *35*, 231–248.

40. Davidson, J.W.; Good, J.M. Social and musical co-ordination between members of a string quartet: An exploratory study. *Psychol. Music* **2002**, *30*, 186–201.
41. Kawase, S.; Obata, S. Audience gaze while appreciating a multipart musical performance. *Conscious. Cogn.* **2016**, *46*, 15–26.
42. Juslin, P.N. Cue utilization in communication of emotion in music performance: Relating performance to perception. *J. Exp. Psychol. Hum. Percept. Perform.* **2000**, *26*, 1797–1813.
43. Malhotra, V.A. The social accomplishment of music in a symphony orchestra: A phenomenological analysis. *Qual. Sociol.* **1981**, *4*, 102–125.
44. Varni, G.; Mancini, M.; Fadiga, L.; Camurri, A.; Volpe, G. The change matters! Measuring the effect of changing the leader in joint music performances. *IEEE Trans. Affect. Comput.* **2019**, *1*, 1–13.
45. Rudolf, M.; Stern, M. *The Grammar of Conducting: A Comprehensive Guide to Baton Technique and Interpretation*, 3rd ed.; Schirmer Books; Maxwell Macmillan Canada; Maxwell Macmillan International: New York, NY, USA, 1994.
46. Green, E.A.H.; Gibson, M.; Malko, N. *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*, 7th ed.; Pearson Prentice Hall: Upper Saddle River, N.J, USA, 2004.
47. Boyes Braem, P.; Braem, T. Expressive gestures used by classical orchestra conductors. *The Symposium on the Semantics and Pragmatics of Everyday Gestures*; Weidler: Berlin, Germany, 2004; pp. 127–143.
48. Veronesi, D. “The guy directing traffic”: Gestures in conducted improvised music between metaphor and metonymy. In Proceedings of the RAaM Workshop *Metaphor, Metonymy & Multimodality*, Amsterdam, Netherlands, 4–5 June 2009.
49. Boyes Braem, P.; Braem, T. A pilot study of the expressive gestures used by classical orchestra conductors. In *The Signs of Language Revisited*; Emmorey, K., Lane, H., Eds.; Psychology Press: New York, NY, USA, 2000; pp. 143–167.
50. Price, H.E.; Winter, S. Effect of strict and expressive conducting on performances and opinions of eighth-grade students. *J. Band Res.* **1991**, *27*, 30–43.
51. Mayne, R.G. An investigation of the Use of Facial Expression in Conjunction with Musical Conducting Gestures and Their Interpretation by Instrumental Performers. Ph.D Thesis, The Ohio State University, Columbus, OH, USA, 1992.
52. Wöllner, C. Which part of the conductor’s body conveys most expressive information? A spatial occlusion approach. *Musicae Sci.* **2008**, *12*, 249–272.
53. Ludwa, C. Assessing the Leadership Potential of Choral Conductors, Ph.D. Thesis, Indiana University, Bloomington, IN, USA, 2012.
54. Fink, L.K.; Lange, E.B.; Groner, R. The application of eye-tracking in music research. *J. Eye Mov. Res.* **2018**, *11*, 1–4.
55. Silvey, B.A. Strategies for improving rehearsal technique: Using research findings to promote better rehearsals. *Update Appl. Res. Music Educ.* **2014**, *32*, 11–17.
56. Poggi, I. Signals of intensification and attenuation in orchestra and choir conduction. *Normas* **2017**, *7*, 33.
57. Poggi, I.; Ansani, A. Forte, piano, crescendo, diminuendo: Gestures of intensity in orchestra and choir conduction. In *Proceedings of the 4th European and 7th Nordic Symposium on Multimodal Communication (MMSYM 2016), Copenhagen, 29–30 September 2016*; Linköping University Electronic Press: Linköping, Sweden, 2017; pp. 111–119.
58. Poggi, I.; D’Errico, F.; Ansani, A. The conductor’s intensity gestures. **2020**. Manuscript submitted for publication.
59. Poggi, I.; Ansani, A. The lexicon of the Conductor’s gaze. *The 5th International Conference on Movement and Computing—MOCO ’18*; ACM Press: Genoa, Italy, June 2018; pp. 1–8.
60. Poggi, I. Lo sguardo del maestro. In *Cultura Popolare, Religione Diffusa, Analisi Qualitativa: Un Sociologo Italiano a cavallo tra due Secoli. Studi in onore di Roberto Cipriani*; Corradi, C., Ed.; Morlacchi: Perugia, Italy, 2018; pp. 233–252.
61. Hall, J.A. *Non-Verbal Sex Differences: Communication, Accuracy and Expressive Style*; Johns Hopkins University Press: Baltimore, MD, USA, 1984.
62. Hall, J.A.; Carter, J.D.; Horgan, T.G. Gender differences in nonverbal communication of emotion. In *Studies in Emotion and Social Interaction. Second Series. Gender and Emotion: Social Psychological Perspectives*; Fischer, A.H., Ed.; Cambridge University Press: Cambridge, UK, 2000; pp. 97–117.

63. Baron-Cohen, S.; Wheelwright, S.; Hill, J.; Raste, Y.; Plumb, I. The “Reading the Mind in the Eyes” test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *J. Child Psychol. Psychiatry Allied Discip.* **2001**, *42*, 241–251.
64. Alaerts, K.; Nackaerts, E.; Meyns, P.; Swinnen, S.P.; Wenderoth, N. Action and emotion recognition from point light displays: An investigation of gender differences. *PLoS ONE* **2011**, *6*, e20989, doi:10.1371/journal.pone.0020989.
65. Maes, P.J.; Amelynck, D.; Lesaffre, M.; Leman, M.; Arvind, D.K. The “Conducting Master”: An interactive, real-time gesture monitoring system based on spatiotemporal motion templates. *Int. J. Hum.-Comput. Interact.* **2013**, *29*, 471–487.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).