The Sound of Teaching Music:

Transmitting expressive skills in piano performance

Atsuko Tominaga

Central European University

Department of Cognitive Science

In partial fulfilment of the requirements for the degree of Doctor of Philosophy

in Cognitive Science

Primary Supervisor: Natalie Sebanz

Secondary Supervisor: Günther Knoblich

Vienna, Austria

January 2023

Declaration of Authorship

I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or which have been accepted for the award of any other degree or diploma at Central European University or any other educational institution, except where due acknowledgment is made in the form of bibliographical reference.

The present thesis includes work that appears in the following papers:

Tominaga, A., Knoblich, G., & Sebanz, N. (2022). Expert pianists make specific exaggerations for teaching. *Scientific Reports, 12*(1), 1-12. https://doi.org/10.1038/s41598-022-25711-3

Tominaga, A., Knoblich, G., & Sebanz, N. (2022). What makes musicians infer teaching intentions? PsyArXiv. https://doi.org/10.31234/osf.io/jtuv4

冨永敦子

Atsuko Tominaga

Acknowledgements

I could not have undertaken this journey without the support of my supervisors, Natalie Sebanz and Günther Knoblich, who have provided me with invaluable feedback and constant guidance throughout my PhD. Thank you for always being patient and supportive.

Many thanks to the members of the SOMBY lab. In particular, I am grateful to our lab managers: Dávid Csűrös, Fanni Takátsy and Candasch Acar for their help with data collection. I also thank my colleagues for their insightful comments and suggestions on my research and for the fun we had over drinks and/or board games.

I would like to thank the CEU Cognitive Science Department, especially Györgyné Finta (Réka), Andrea Jenei, Eszter Salamon and Anita Apor for their administrative assistance. Thank you for all your kindness and support.

My gratitude extends to Martin Clayton for introducing me to the world of ethnomusicology and supervising my qualitative study in Durham. I would like to thank my interviewees, who voluntarily took part in my study and dedicated time to share their personal experiences about teaching and learning music.

Special thanks to my online friends on fediverse. Thank you for giving me useful computer/programming-related advice and teaching me useless internet memes. Without you guys, my PhD would have been very lonely and depressing. Thank you for always being with me while I was suffering and struggling in front of my laptop.

Ljuba, thank you for listening to me all the time and talking with me about research, academia, relationships, life and so on. I have learnt a lot from you as a woman.

Richard, thank you for your encouragement and emotional support during the hardest time of my PhD. Thank you for being with me and growing up together.

My warm and sincere thanks go to my parents and brother for their tremendous support at every stage of my life.

Lastly, Budapest and Vienna – Thank you for giving me an opportunity to have such a wonderful journey in Central Europe, the place I have always dreamed of.

Abstract

Social learning plays a pivotal role in human skill acquisition and is an integral part of cultural transmission. This dissertation aims to investigate the role of teaching in transmitting expressive musical skills that require learning *how* to perform actions to implement particular expressive techniques. In the first chapter, I discuss how people use movement kinematics to convey pedagogical intentions and how these intentions may be detected by observers. I then report four empirical studies. Our first study examined whether and how expert pianists modulate their playing when they have the intention to teach musical expressive techniques such as articulation (smoothness of sound) and dynamics (loudness of sound) compared to when they perform to an audience. We found that expert pianists exaggerated relevant aspects of the techniques to be taught (e.g., they exaggerated dynamics contrasts between forte and piano when teaching dynamics). Our second study investigated whether and how expert pianists adapt their playing depending on novices' demonstrated skills. The results showed that expert pianists highlighted articulation-related aspects when novices did not implement articulation as notated. Our third study investigated what makes musician listeners infer pedagogical intentions based on listening to recorded performances. We demonstrated that musician listeners were sensitive to cues distinguishing recordings produced for teaching from those produced for performing to an audience. Performances with slower tempo and relevant exaggerations (e.g., exaggerated legato and staccato when performing the piece with notated articulation) were likely to be considered as performed for teaching purposes. Lastly, our fourth study explored qualitative aspects of expressive performance by conducting semi-structured interviews with music teachers and students. In the last chapter of the dissertation, I discuss the implications and the limitations of our approach by considering ostensive communication in skill transmission, interactivity between experts and novices, and the goal of expressive performance in music. The findings of this dissertation reveal the importance of action-based communication between experts and novices to transmit expressive skills in piano performance.

Table of Contents

Chapter 1. Introduction

Human skill transmission is achieved through social learning between experts and novices (Tomasello, Kruger, & Ratner, 1993). Beyond passing on generic knowledge, humans have developed expertise in various skills from stone knapping to making music that they pass on (Ericsson, 2008; Mesoudi & Thornton, 2018; Savage, 2019; Stout, 2011). Expertise involves the ability to perform highly complex actions and in artistic contexts also requires expressive performance. For instance, expert pianists are skilled with regard to not only bimanual motor coordination but also rich expressions and interpretations of music. Expressive skills are one of the important elements of artistic creativity and can hardly be acquired by novices without deliberate practice and experts' support (Lehmann & Ericsson, 1997; Sloboda, 2000). In this dissertation, I will explore the role of teaching in musical skill transmission, focusing particularly on expressive performance.

1.1 Teaching as joint action

Teaching is one of the essential components of social learning (Tennie, Call, & Tomasello, 2009; Whiten, 2017). Particularly when it comes to acquiring expertise, domain-specific knowledge and skills are not merely learnt but explicitly taught by knowledgeable experts (Ericsson, 2008; Heyes, 2017). Teaching is considered to facilitate learning in novices by directly communicating relevant information to them (Kline, 2015; Thornton & Raihani, 2008). Teaching may shift learners' attention and affect cognitive processes necessary for learning, such as memory and imitation (Csibra & Gergely, 2009; Király, Csibra, & Gergely, 2013; Okumura, Kanakogi, Kobayashi, & Itakura, 2020; Yoon, Johnson, & Csibra, 2008).

Demonstration is considered to be one of the fundamental ways of teaching especially when experts can directly communicate with novices and interact dynamically in time and space (Kline, 2015; Strauss & Ziv, 2012). Experts first demonstrate what novices need to acquire while novices are asked to observe and then copy what the experts demonstrated. It is rare that experts demonstrate a whole sequence of complex actions at once, but rather they decompose the complex actions into smaller units and incrementally add complexity (Marchand, 2010). Novices start discovering hierarchically complex structures of the actions they need to acquire by parsing the actions through imitative learning (Byrne, 2003). Evidence from dance expertise acquisition showed that learning from experts' decomposition of movements is more effective than learning from novices' own decompositions, especially for introductory dancers (Rivière, Alaoui, Caramiaux, & Mackay, 2019).

Experts use various ways of communication during demonstrations. The use of language is common and considered to be effective in skill transmission (Lombao, Guardiola, & Mosquera, 2017; Morgan et al., 2015); however, spoken language coincides with non-verbal bodily cues such as deictic pointing, posture, gesture and facial expression (Marchand, 2010; Moro, Mortimer, & Tiberghien, 2020). Marchand (2010) reported how a carpentry teacher used different channels of communication to achieve different goals. For example, the teacher used language to deliver concepts and ideas of skill and to promote interactive learning (e.g., questioning). While he was talking, he modulated the way of speaking by changing tone and volume to emphasise particular aspects of information. While explaining the skill verbally, the teacher spent extensive time demonstrating skilled actions physically to show how to use a specific tool. The teacher also monitored how students acted and provided help if necessary (e.g., he intercepted students' manual actions and demonstrated correct procedures).

Teaching is a dynamic process, which can be achieved through the mutual contribution of experts and novices (De Felice, Hamilton, Ponari, & Vigliocco, 2023; Watanabe, 2013). One might think that the information flow between experts and novices is unidirectional (i.e., from expert to novice). However, novices are not merely able to copy and replicate what experts taught, but rather explore and develop knowledge and skills on their own. As I showed in the example of the carpentry teacher, knowledge and skills are firstly deconstructed by experts and novices have to reconstruct them by observing and imitating the experts. At the beginning of the learning process, the important role of experts is to monitor novices' actions, identify errors and provide appropriate feedback so as to improve novices' performance (Ericsson, 2008). With the help of experts, novices become able to monitor their own actions and detect errors by themselves. Marchand (2010) observed that students sometimes complemented what the teacher was going to say once they reached a certain skill level and that it looked as if the role of the teacher and the student was swapped. In other words, it can be said that students became able to predict what the teacher is going to say or do. Thus, teaching can be seen as one form of joint action where two or more people act together in time and space (Sebanz, Bekkering, & Knoblich, 2006). By engaging in joint action with experts, novices seem to simulate others' actions using their own bodies and develop cognitive skills such as monitoring and predicting their own and others' actions (Wolpert, Doya, & Kawato, 2003).

In the following section, I will discuss how experts send pedagogical signals for novices to deliver relevant information during demonstration and how they adapt their behaviour depending on the knowledge and skills of novices.

1.2 Sending pedagogical signals during demonstration

Sensorimotor communication is employed to provide task-relevant information to co-actors without the use of language while engaging in joint action (Sebanz & Knoblich, 2021). Numerous studies about sensorimotor communication have shown that people use subtle bodily movements to convey intentions to others (Pezzulo, Donnarumma, & Dindo, 2013; Pezzulo et al., 2019). Sensorimotor communication refers to communicative signals that are embedded in a pragmatic action. For example, when you want to move a table together with your partner, you may push the table slightly to a certain direction to move the table (a pragmatic action) and also to inform your partner about the direction you want to move towards (a communicative action). Another example would be when violinists are performing the same piece together as an ensemble, they may exaggerate their arm movements to make a sound (a pragmatic action) and also to coordinate the sound onset with each other (a communicative action).

Sensorimotor communication has also been observed in teaching contexts. Research on infant-directed speech and action has shown that caregivers are likely to exaggerate their speech and action when they have the intention to teach or demonstrate skills to others. For instance, caregivers are likely to exaggerate their speech by producing higher pitch or larger pitch contours to help infants process complex language structures (so-called motherese; Fernald & Kuhl, 1987; Saint-Georges et al., 2013). Likewise, when caregivers are demonstrating how to use a novel toy to infant learners, they tend to exaggerate their movements to help infants process complex action sequences (so-called motionese; Brand et al., 2002). Similar teaching behaviour has been found in individuals teaching adult learners. For example, native British English speakers used some properties of motherese when talking to adult learners who were studying English, and skilled adults exaggerated their movements when teaching a novel xylophone melody to adult novices (McEllin, Knoblich, & Sebanz, 2017; Uther, Knoll, & Burnham, 2007). These action modulations have the function to highlight and deliver relevant information to learners (Vesper & Sevdalis, 2020).

One of the pivotal aspects of teaching is that experts monitor novices' performance and flexibly adapt their intervention for each individual at each level of skill acquisition (Byrne & Rapaport, 2011; Ericsson, 2008). Generally, in asymmetric joint action where one partner has full task information (i.e., more knowledgeable interaction partner) whereas the other partner has partial task information (i.e., less knowledgeable interaction partner), it has been shown that more knowledgeable interaction partners tend to modulate their behaviour depending on what their less knowledgeable interaction partners know. For example, people considered the past history of interactions with their partners who have less task information when sending communicative signals and stopped sending the signals when their partners had full task information (Candidi, Curioni, Donnarumma, Sacheli, & Pezzulo, 2015; Pezzulo & Dindo, 2011). In teaching contexts, experts seem to monitor novices' performance and adapt their demonstration to the abilities of novices. For instance, when caregivers demonstrated how to stack and fit four nesting cups to their infants, they dynamically modulated their demonstration depending on the infants' knowledge and skills (Fukuyama et al., 2015). The caregivers stopped exaggerating their movements when either the infants succeeded at reproducing the expected actions or when it was obvious that the infants did not have the motor skills necessary to produce the actions. Therefore, pedagogical signalling by experts is adaptive to novices' demonstrated skills.

In this section, I outlined how experts act in order to send pedagogical signals to novices and how such signalling is flexible depending on the progress of novices' skill acquisition. For successful communication in skill transmission, it is crucial that pedagogical cues produced by experts can be detected by novices and can be used for learning. In the next section, I will therefore review how observers (potential learners) recognise and infer pedagogical intentions by observing demonstrations by experts.

1.3 Inferring pedagogical intentions from demonstration

People have the propensity to infer intentions from observed actions performed by others (Blakemore & Decety, 2001). Observers can decode intentions from observed actions and distinguish different social intentions such as whether actors have the intention to cooperate or compete (Manera, Becchio, Cavallo, Sartori, & Castiello, 2011; Sartori, Becchio, & Castiello, 2011). Moreover, observers seem to be able to infer intentions by focusing on particular movement kinematics, not an entire movement pattern. For example, participants tended to judge that observed bottle-grasping movements were produced in order to pour water into a glass when wrist height was relatively lower whereas they tended to consider that the grasping movements were produced in order to drink water when wrist height was relatively higher (Cavallo, Koul, Ansuini, Capozzi, & Becchio, 2016). These studies indicate that subtle body movements might be sufficient to deliver what actors intend to do to observers.

Perceiving pedagogical intentions from experts' actions is one of the fundamental components of social learning (Csibra & Gergely, 2009). In teaching contexts, experts' behavioural modifications seem to attract novices' attention. For example, infants prefer infant-directed speech and action (i.e., motherese, motionese) over adult-directed speech and action (Brand & Shallcross, 2008; Fernald, 1985). Given that motherese and motionese seem to positively affect infants' learning processes (Koterba & Iverson, 2009; Saint-Georges et al., 2013), the possibility can be considered that infants recognise and make use of pedagogical cues produced by adults. For adult learners, McEllin, Sebanz and Knoblich

(2018) examined whether people could discriminate between actions performed with the intention to coordinate (i.e., performing a xylophone melody together with a partner) and actions performed in order to teach (i.e., performing the melody for a novice partner) by observing the movement kinematics of performers playing the xylophone. They found that observers could distinguish the intention to coordinate from the intention to teach by focusing on specific movement parameters (e.g., velocity and movement amplitude). Therefore, there seem to be specific characteristics of actions (e.g, slower or exaggerated performance) that make observers infer pedagogical intentions.

The ability to perceive intentions from observed actions is influenced by the expertise of observers. When playing basketball, a player might produce their movements not only to pass the ball to his/her teammate but also to deceive his/her opponent's team. Sebanz and Shiffrar (2009) demonstrated that experts were better at predicting if a player intended to fake or make a pass than novices when they observed the player’s action. Moreover, Aglioti, Cesari, Romani and Urgesi (2008) examined how people who have different expertise anticipate the success of free shots in basketball (i.e., experts (athletes), expert watchers (coaches or sports journalists), and novices). They found that experts were better at predicting the outcome of free shots at the early stage of observation compared to expert watchers and novices. It was proposed that experts use their own motor system to simulate and predict the success of free shots whereas expert watchers and novices seemed to focus on the trajectory of the shots. Therefore, in order to perceive subtle bodily cues, motor expertise might be a prerequisite to infer intentions when it comes to observing skilled performance. This means that while novices may become aware of teaching intentions based on kinematics, subtle predictions of observed movements may be outside of their expertise.

In this section, I discussed whether pedagogical cues produced by experts can be perceived and potentially used by observers (potential learners) for skill acquisition. So far, I have introduced cases of normative performance, where any deviation from the performance (e.g., slower or exaggerated performance) is considered to signal informative intentions such as highlighting relevant information for learners. In the next section, I will consider how sensorimotor communication can be used for pedagogical purposes in the context of teaching expressive performance, where the performance itself necessarily deviates from normative performance.

1.4 Teaching expressive performance in music

Another function of sensorimotor communication is to elicit emotional and aesthetic experiences in an audience (Vesper & Sevdalis, 2020). In artistic contexts, performers intentionally deviate from normative performance and try to produce expressive performance to express their emotions and interpretations of art. In music, expressivity is considered the most important performance factor (Laukka, 2004; Sloboda, 2000). Musicians use various expressive tools such as tempo, articulation (smoothness), dynamics (loudness), timbre as well as ancillary gestures to convey intentions and interpretations of music to an audience (Meissner, 2021). Meissner (2021) raised the possibility that children's experiences with infant-directed speech and song may help the development of communication and expressive skills in music performance. However, most studies in music psychology seem to assume that verbal teaching (e.g., verbal approval/disapproval, explicit instructions about musical properties, using metaphor and mental imagery) or aural modelling (i.e., learners listening to and imitating the performances of models (e.g., teachers, professional musicians or recordings)) is effective (Speer, 1994; Woody, 2000, 2006). Little research has focused on how real-time interactions through sensorimotor communication between music teachers and students might be beneficial for acquiring expressive performance (But see; Li & Timmers, 2020, 2021).

A potential problem with the use of sensorimotor communication when teaching expressive performance is that it is not straightforward how experts should deviate from normative performances for teaching purposes. According to the findings in the literature of motherese and motionese, it can be expected that experts may perform slowly and exaggerate relevant aspects of skills to be taught. However, these performance modulations may then provide another interpretation of the music. In order to investigate whether and how music experts modulate their expressive performance for teaching, I performed a study with two experiments where expert pianists were asked to perform one piece of music with expressive notations (articulation or dynamics). In one condition, I asked participants to perform the piece with a designated expressive notation with the intention to teach it to their students whereas, in the other condition, I asked participants to perform the piece with the intention to perform it to an audience. This study aims to answer whether and how expert pianists modulate their performance for teaching compared to an expressive performance baseline (Chapter 2). In a further study, I investigated how expert pianists' performance modulations change according to novices' demonstrated skills (Chapter 3).

In addition to focusing on the production (Chapter 2) and the adaptation (Chapter 3) of experts' expressive performance in teaching contexts, I performed a study addressing novices' perception of experts' expressive performance for teaching (Chapter 4). It is not clear whether and how observers perceive and infer teaching intentions when listening to expressive performance. As Jacob and Jeannerod (2005) pointed out, it is unlikely that people accurately recognise communicative intentions produced by actions without additional information when there are more than two different possible goals of the actions. In the case of expressive performance in teaching contexts, it may be challenging for observers to judge whether a performance was produced for expressing an interpretation of music or for teaching the expressive aspect of the performance. In order to investigate which performance modulations make listeners infer teaching intentions, I asked participants to listen to a number of recordings taken from the study reported in Chapter 2 and to judge if each recording had been produced for teaching purposes or not. To this end, we correlated participants' judgments for teaching with quantified performance parameters such as tempo, smoothness and loudness.

In Chapter 5, I explore creative aspects of expressive performance, going beyond the scope of my experimental studies by conducting semi-structured interviews with music teachers and students. Some music researchers have argued against the idea that knowledge and skills can be conceptualised independently from the context and have questioned whether transmission happens unidirectionally from experts to novices (Borgo, 2007, Van der Schyff, Schiavio, & Elliott, 2016). Rather, they claim that learning happens during dynamic interactions between experts and novices and cannot be separated from the context. Rice (2003) argued that informal music learning by participating in social gatherings seems to be more common than learning music through formalised music training with teachers at school or in a private lesson. Therefore, the teaching model I proposed here (*Fig 1.1*) does not seem to capture the whole process of teaching and learning in the real world. Through dialogues with music teachers and students, I investigate qualitative aspects of expressive performance by focusing on the individual process of becoming a musician. Chapter 6 provides a discussion of the implications and limitations of our approach. Particularly, I consider how ostensive communication and interactive learning with experts contribute to skill acquisition and discuss the goal of expressive performance in music.

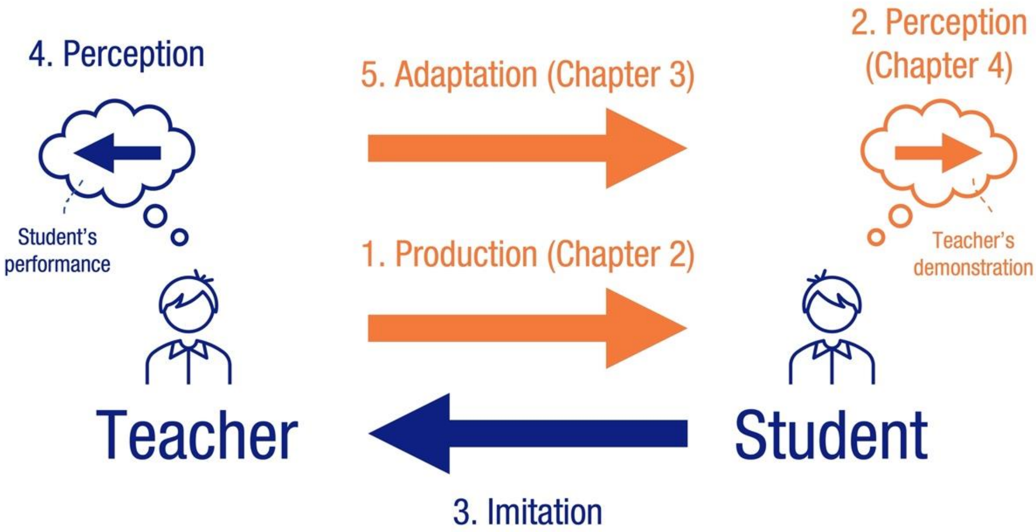

Figure 1.1. Schematic representation of the interactive skill transmission process investigated in this dissertation. Particularly, I focused on the three aspects of the process highlighted in orange colour.

Chapter 2. Sound exaggerations for teaching

2.1 Introduction

Deliberate teaching has supported human skill transmission over generations and provides a key route for learning (Thornton & Raihani, 2008; Tomasello, 2016). Experts modulate their behaviour so that novices can extract relevant information to learn a novel skill. For instance, adults modulate their speech (motherese) and actions (motionese) when demonstrating a skill to infant learners (Brand et al., 2002; Saint-Georges et al., 2013). These modulations include slowing down and exaggerating sounds and actions. Similar findings were obtained in studies with adult learners, where exaggerations were observed when native British English speakers were talking to second language English learners (Uther et al., 2007) and when skilled adults were teaching xylophone melodies to novices (McEllin et al., 2017). The observed modulations are thought to provide communicative signals that can facilitate learning by affecting attention and memory (Csibra & Gergely, 2009).

Previous research has shown that demonstrators not only adjust their gestures (Campisi & Özyürek, 2013) and actions (Fukuyama et al., 2015) to learners' skills but engage in specific action modulations to highlight certain aspects of demonstrated actions. For example, Schaik and colleagues (2019) showed that adults used specific action modulations for demonstrating different action effects of objects to infants. Ho and colleagues (2016) found that demonstrators chose costly movement paths at structurally important points to disambiguate one goal over other possibilities.

Performing actions with exaggeration is straightforward for actions that are normally performed in the most efficient way possible (Pezzulo et al., 2013). However, how can particular aspects of actions be highlighted when the actions themselves are expressive even outside of a teaching context? This is the case in music performance, where pieces are played with expression. Expressivity is a vital component of performance and typically the main focus of music teaching (Laukka, 2004). Expressive skills are generally considered to be separate from technical skills, however, fine motor control is required to implement subtle sound modulations in expressive performance (Sloboda, 2000). In some music genres such as Western classical music, it is crucial to acquire the motor skills needed to perform a piece expressively. For example, pianists translate their own interpretations of music and convey their emotions by modulating specific parameters such as timing, smoothness and loudness of sound. This raises the question of whether and how musicians modulate their actions during pedagogical demonstration of expressive techniques.

In naturalistic teaching settings, teachers have many possibilities for how to convey to a learner how to play a piece expressively. They may use verbal communication (Woody, 2000, 2006) and non-verbal bodily cues (Simones, Schroeder, & Rodger, 2015) as well as modulating the sounds that they are producing. In the present study, we focused entirely on the teaching of expressive techniques through sound in order to determine whether expert musicians would systematically exaggerate the actions necessary to implement a particular technique for musical expression. Pianists were asked to perform a piece to demonstrate to a learner how to implement the notated expressions (teaching condition) and to play the same piece for an audience (performing condition). Two basic expressive techniques in piano performance were used: articulation (legato and staccato) and dynamics (forte and piano). Articulation was assessed by key-overlap time between two consecutive notes. Positive key-overlap time indicates legato styles whereas negative key-overlap time indicates staccato styles. Dynamics was evaluated by key velocity of each key press. Higher key velocity means forte styles while lower key velocity means piano styles.

If experts rely on generic action modulations, it can be expected that they will play more slowly during pedagogical demonstration, regardless of the kind of techniques to be taught. To the extent that experts use action modulations to support their teaching of specific techniques, they should exaggerate legato (i.e., key-overlap time should be more positive) and staccato (i.e., key-overlap time should be more negative) when teaching articulation whereas they should exaggerate forte (i.e., key velocity should be higher) and piano (i.e., key velocity should be lower) when teaching dynamics. Furthermore, one could speculate that they might produce modulations specifically at structurally important points that best highlight the technique to be taught. Importantly, they should avoid modulating irrelevant properties of expression (e.g., the smoothness of sound while teaching dynamics).

In Experiment 1, we employed a simple musical scale to examine whether and how skilled pianists vary their performance depending on which expressive techniques (i.e., either articulation or dynamics) they are teaching. Experiment 2 was conducted to replicate our findings from Experiment 1 with a more naturalistic piece of music.

2.2 Experiment 1

2.2.1 Methods

Participants. We recruited 36 piano experts who played the piano for at least the past 10 years or were studying advanced piano performance at a music school at the time of recruitment. For data analysis, we excluded three participants due to experimental errors, and two participants because they deviated substantially from the prescribed tempo (outside 2 standard deviations from the average tempo across participants). Thirty-one participants (15 female) were included in data analysis. Most participants were right-handed (left: 2, ambidextrous: 2) with a mean age of 24.16 (*SD* = 4.26). They had 12.45 years of practice on average (*SD* = 5.66). 10 participants had experience in teaching the piano (*M* = 3.48 years, *SD* = 3.51). All participants gave their informed consent before the experiment started and received vouchers for their participation. The study (No. 2018-124) was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary.

Apparatus and stimuli. A weighted Yamaha MIDI digital piano was used to record participants' performance via Max/MSP (https://cycling74.com/products/max) on a MacBook Pro with Mac OS X Mojave 10.14.3. The laptop and piano were connected to a high-fidelity soundcard (Focusrite Scarlett 6i6) to deliver a metronome and piano sound. All auditory feedback was given to participants through headphones (Audio-Technica ATH-M50X). Sheet music was displayed on a computer monitor in front of the participants. The pitch, onset and offset time of each note, and key velocity profiles were obtained from MIDI data using Max/MSP patchers.

One musical excerpt was used as a stimulus. The excerpt was taken from "A Dozen a Day - Play with Ease in Many Keys" by Edna-Mae Burnam and modified for the experiment. It consisted of a 6-measure isochronous melody noted in a 4/4 metre. The stimulus was composed in C major to be played with the right hand only. Original sheet music (i.e., sheet music without expressive notations, *Fig 2.1 A*) was used for the purpose of practice. Expressive notations were added to the original sheet music for the experiment. They referred to either articulation or dynamics (*Fig 2.1 B-C*). Articulation was notated as either legato or staccato. Legato indicates that musical notes are to be connected and should sound smooth. Staccato requires producing musical notes with shortened duration, keeping them separate from each other. Dynamics was notated as either forte or piano. Forte indicates that musical notes should be played loudly whereas piano indicates that musical notes should be played softly. The notation did not include any indication of fingering (i.e., the positioning of the fingers when playing the piano) because the piece was simple and pilot testing had shown that specifying fingering was not necessary.

Procedure. First, participants were allocated to either the teaching or performing condition and asked to practise the excerpt with the notated expression of either articulation or dynamics (*Fig 2.1 B-C*). As soon as they had produced the excerpt with the notated expression without pitch errors twice consecutively, the test trials began[1].

In the teaching condition, participants were instructed to play the excerpt of music as if they were teaching it to students. It was mentioned that the students already knew the sequence of the tones and that they were trying to learn how to perform the piece with the notated expression by listening to the participant's performance. In the performing condition, participants were asked to play the excerpt of music as if they were performing it to an audience (see details in *Supplementary Material 1*). Participants played the piece 8 times per technique per condition, so there were 32 trials in total (2 conditions x 2 techniques x 8 trials). The order of the conditions was blocked and counterbalanced across participants. The order of the techniques within each condition was also blocked and counterbalanced across participants. A leading metronome (80 quarter beats per minute, 8 beats) indicated the target tempo before each trial.

At the end of the experiment, participants filled in a questionnaire asking about their demographic information and experience in piano performance/teaching.

Data analysis. Three dependent variables were computed for data analysis. Inter-onset intervals (IOIs) are the intervals between onsets of adjacent notes and provide a

[1] Participants were not allowed to use the pedal throughout the experiment.

measure of tempo. Key-overlap time (KOT) is the difference between the offset time of the current tone (i.e., key release time) and the onset time of the ensuing tone and is a measure for the smoothness of musical sequences (Bresin & Battel, 2000). A positive value indicates smooth legato styles due to overlap between the current and ensuing tone whereas a negative value indicates sharp staccato styles due to separation between the current and ensuing tone. Tone intensity is assessed by key velocity (KV) and measures the loudness of a musical note. A higher value indicates forte styles whereas a lower value indicates piano styles. The value of KV in MIDI varies between 0 (minimum) and 127 (maximum).

Data cleaning, preprocessing and statistical analysis were performed in R version 4.0.5. For statistical analysis, only 16th notes with expressive notations were included. Overall, five trials were excluded from data analysis because participants did not follow the sheet music or stopped performing before the end. Pitch errors were identified by comparing the sequence of musical notes produced by a participant with the sequence of musical notes according to the sheet music. Pitch errors included either extra, missing or substituted tones and were manually removed by using the *editData* R package. For onsets, 11.65% of the trials contained at least one pitch error (extra notes: 6.28%, missing notes: 5.07%, substituted notes: 0.30%). For offsets, 14.08% of the trials contained at least one pitch error (extra notes: 6.28%, missing notes: 5.17%, substituted notes: 2.63%). We found that some participants did not precisely follow the sheet music (e.g., they held some notes longer than notated), therefore the order of offsets did not correspond to that of onsets. We counted these as errors and removed the erroneous notes even if the order of onsets was correct. As a result, less than 1 % of total responses were corrected. In addition to pitch errors, we removed outliers for IOIs, KOT and KV, defined as values more than 3 standard deviations from the mean of each dependent variable. For each dependent variable, this resulted in less than 5% of overall responses being removed as outliers.

We performed separate analyses for the two techniques (i.e., articulation and dynamics). A paired-sample *t*-test or a Wilcoxon Signed-rank test (a non-parametric alternative to a paired *t*-test) was performed to compare the mean IOIs in the teaching and performing condition. For KOT and KV, we performed a 2 x 2 repeated-measures analysis of variance (ANOVA) with the factors Condition (teaching vs. performing) and Subcomponent (Articulation: legato vs. staccato or Dynamics: forte vs. piano, respectively). The *t.test* or *wilcox.test* function in the *stats* R package was used for a *t*-test or a Wilcoxon Signed-rank test. For calculating an effect size, we used the *rstatix* R package. The *aov_car* function in the *afex* R package was used for a repeated-measures ANOVA. For post-hoc comparisons on the estimated marginal means, we used the *emmeans* R package.

Figure 2.1. (A)Original sheet music. (B)Articulation. The curved line (slur) indicates legato and the dots indicate staccato. (C)Dynamics. The symbol 'f' denotes forte and the symbol 'p' denotes piano. For data analysis, only the 16th notes were included.

2.2.2 Results

All effects are reported as significant at $p < .05$. We first report results for performance of the piece with the notated expression of articulation (*Fig 2.1 B*), followed by performance of the piece with the notated expression of dynamics (*Fig 2.1 C*). To recall our predictions, if participants play more slowly when they are trying to teach, inter-onset intervals (IOIs) should be larger when teaching. If participants specifically modulate relevant aspects of the techniques they are trying to teach, key-overlap time (KOT) should be more positive for legato and more negative for staccato when teaching articulation, and key velocity (KV) should be higher for forte and lower for piano when teaching dynamics.

Articulation

Inter-onset Intervals (IOIs). To compare the mean IOIs between the teaching and performing condition, we conducted a Wilcoxon Signed-rank test, instead of a paired *t*-test, because a Shapiro-Wilk test showed that the distribution of the mean difference was significantly different from the normal distribution ($p < 0.001$). The Wilcoxon Signed-rank test revealed that participants played more slowly in the teaching condition [*Mdn* = 189.55 (ms), *IQR* = 26.74] than in the performing condition [*Mdn* = 185.23 (ms), *IQR* = 19.30] while playing the piece with the notated articulation ($p = 0.021$, $r = 0.41$, two-tailed, *Fig 2.2*).

Key-Overlap Time (KOT). A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Articulation (legato vs. staccato) revealed that there was a significant main effect of Articulation ($F(1,30) = 1149$, $p < 0.001$, $\eta_G^2 = 0.94$) and a significant interaction between Condition and Articulation ($F(1,30) = 8.52$, $p = 0.007$, $\eta_G^2 = 0.010$, *Fig 2.3*). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment showed that for staccato KOT was more negative in the teaching condition

[*M* = -133.79 (ms), *SD* = 20.13] than in the performing condition [*M* = -128.34 (ms), *SD* = 20.86] (*p* = 0.005). For legato there was no significant difference in KOT between the teaching [*M* = 15.90 (ms), *SD* = 15.51] and performing condition [*M* = 14.01 (ms), *SD* = 17.02] (*p* = 0.32).

Key Velocity (KV). A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Articulation (legato vs. staccato) showed that there was a significant interaction between Condition and Articulation ($F(1,30) = 7.45$, $p = 0.011$, $\eta_G^2 = 0.004$, *Fig 2.4*). However, post-hoc comparisons based on the estimated marginal means with Tukey adjustment did not find a significant difference between the teaching [Legato: *M* = 70.34, *SD* = 6.46; Staccato: *M* = 72.00, *SD* = 8.74] and performing condition [Legato: *M* = 70.68, *SD* = 5.86, Staccato: *M* = 70.41, *SD* = 7.97] for each subcomponent (Legato: *p* = 0.68, Staccato: *p* = 0.09).

Dynamics

Inter-onset Intervals (IOIs). To compare the mean IOIs between the teaching and performing condition, we conducted a Wilcoxon Signed-rank test, instead of a paired *t*-test, because a Shapiro-Wilk test showed that the distribution of the mean difference was significantly different from the normal distribution ($p < 0.001$). The Wilcoxon Signed-rank test revealed no significant difference between the teaching condition [*Mdn* = 186.28 (ms), *IQR* = 21.83] and performing condition [*Mdn* = 186.27 (ms), *IQR* = 19.59] while playing the piece with the notated dynamics (*p* = 0.11, *r* = 0.29, two-tailed, *Fig 2.2*).

Key Velocity (KV). A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Dynamics (forte vs. piano) showed that there was a significant main effect of Dynamics ($F(1,30) = 369$, $p < 0.001$, $\eta_G^2 = 0.74$) and a significant interaction between Condition and Dynamics (*F*(1,30) = 6.83, *p* = 0.014, η_G^2 = 0.003, *Fig 2.4*). However, post-hoc comparisons based on the estimated marginal means with Tukey adjustment did not find a significant difference between the teaching [Forte: *M* = 82.78, *SD* = 8.31, Piano: *M* = 60.63, *SD* = 5.08] and performing condition [Forte: *M* = 81.65, *SD* = 7.43, Piano: *M* = 60.99, *SD* = 4.41] for each subcomponent (Forte: *p* = 0.068, Piano: *p* = 0.24).

KV Transition Points. Additionally, we conducted an exploratory analysis by focusing on specific points at which each subcomponent changed from one to the other (i.e., forte to piano: FtoP, piano to forte: PtoF). These points could be structurally important to make a contrast between forte and piano. We calculated the KV difference for each interval by subtracting the KV value of the current note from that of the following note. Outliers were removed using the same criteria as for the other dependent variables and less than 5 % of the data were excluded from analysis. A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Transition Type (FtoP vs. PtoF) showed that there was a significant main effect of Transition Type (*F*(1,30) = 224, *p* < 0.001, η_G^2 = 0.83) and a significant interaction between Condition and Transition Type (*F*(1,30) = 25.99, *p* < 0.001, η_G^2 = 0.033, *Fig 2.9*). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment showed that there was a larger KV difference when changing from forte to piano (*p* = 0.002) and from piano to forte (*p* < 0.001) in the teaching condition [FtoP: *M* = -14.11, *SD* = 7.17; PtoF: *M* = 24.27, *SD* = 9.77] than in the performing condition [FtoP: *M* = -11.70, *SD* = 5.99, PtoF: *M* = 20.87, *SD* = 8.67].

Key-Overlap Time (KOT). A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Dynamics (forte vs. piano) revealed that there was a significant main effect of Dynamics (*F*(1,30) = 308, *p* < 0.001, η_G^2 = 0.30, *Fig 2.3*), reflecting more key-overlap for forte notes compared to piano notes. Neither the main effect of Condition (F(1,30) = 0.60, p = 0.44, η_G^2 = 0.000) nor the interaction between Condition and Dynamics (F(1,30) = 0.053, p = 0.82, η_G^2 = 0.000) was significant.

2.2.3 Discussion

The findings from Experiment 1 indicated that skilled pianists modified their performance for teaching purposes. For IOIs, we found a small but significant slowing down during teaching specifically when playing the piece with the notated articulation. This finding is in line with earlier studies that found slower performance of actions in a teaching context (e.g., Brand et al., 2002; McEllin et al., 2017). However, we did not find a significant difference in tempo when participants were teaching dynamics. It could be that trying to keep the prescribed tempo (conveyed through the leading metronome) limited the extent to which participants slowed down their performance during teaching. Another possibility is that slower performance may be beneficial to highlight the relation between two notes (i.e., to what extent two notes overlap) and was therefore employed when teaching articulation, whereas slower performance might not help when teaching dynamics.

The results for KOT and KV showed that participants successfully highlighted relevant aspects of articulation and dynamics. Specifically, participants exaggerated staccato when teaching articulation. Moreover, our exploratory analysis demonstrated that participants made a larger contrast in dynamics at transition points (i.e., forte to piano or piano to forte). Importantly, participants did not modulate their performance in terms of irrelevant aspects of the techniques for teaching purposes (e.g., modulating the smoothness of sound while teaching dynamics). These findings confirmed that participants modulated their performance in systematic and fine-grained ways to teach particular techniques.

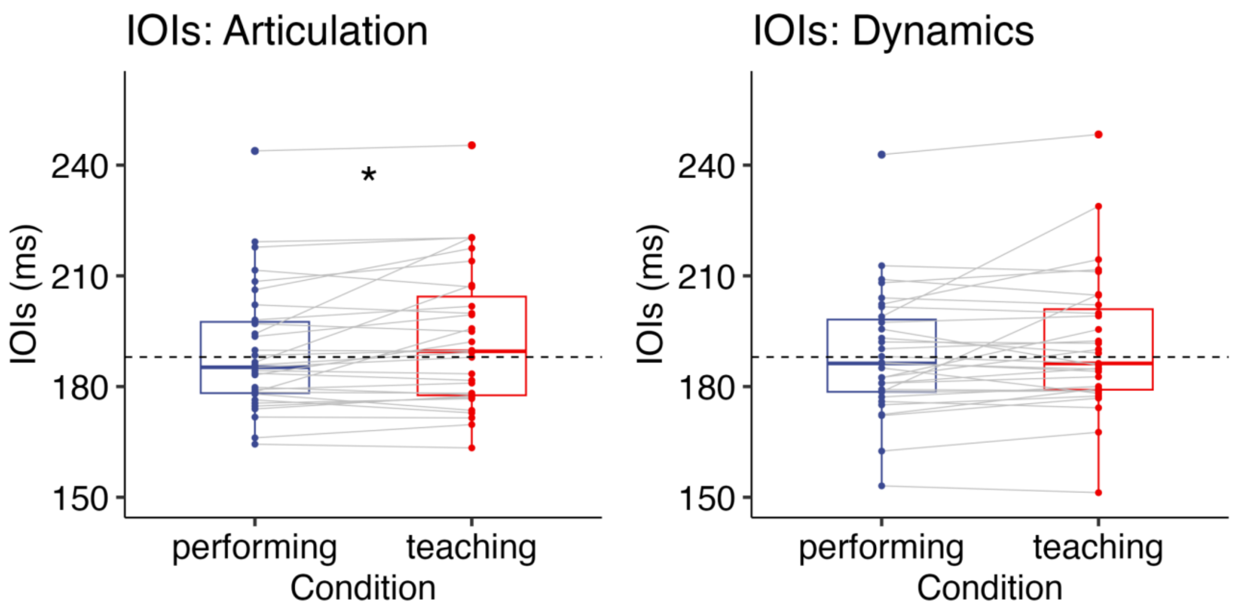

*Figure 2.2. Experiment 1: IOIs (ms) when playing the piece with either articulation (left) or dynamics (right). A dashed line represents the tempo given by a metronome. Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

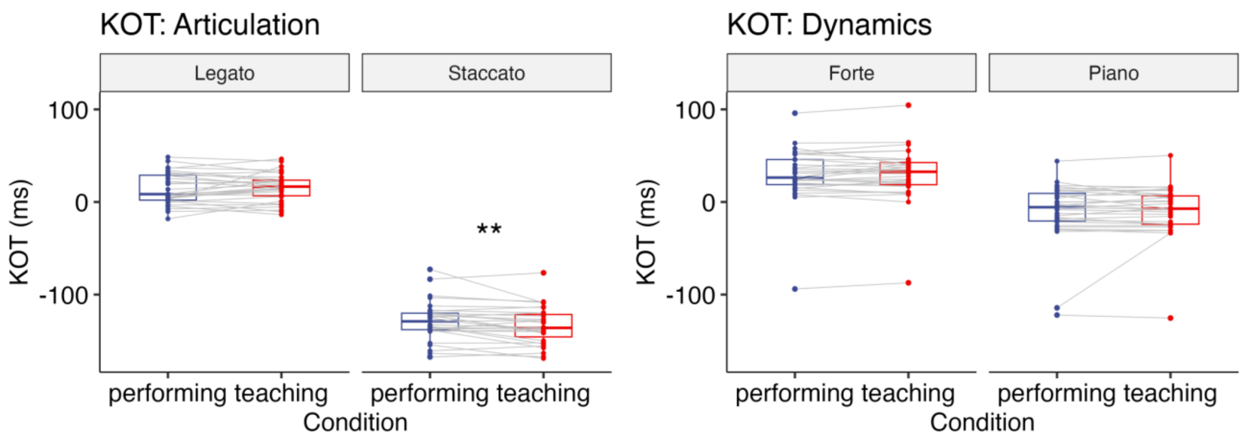

*Figure 2.3. Experiment 1: KOT (ms) when playing the piece with either articulation (left) or dynamics (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

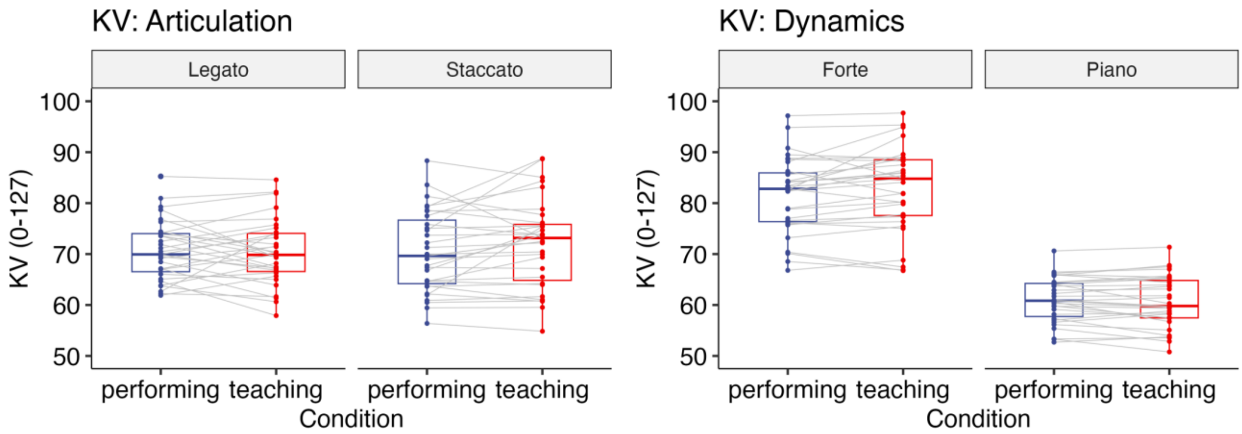

*Figure 2.4. Experiment 1: KV (0-127) when playing the piece with either articulation (left) or dynamics (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

2.3 Experiment 2

In Experiment 1, we employed a simple musical scale to maximise experimental control. To test whether our results generalise to a more naturalistic piece, in Experiment 2, we chose an actual piano piece and modified it for the purpose of the experiment. If pianists selectively highlight the relevant aspects of the techniques to be taught also in a more naturalistic piece containing more opportunities for expression, key-overlap time should be again more positive for legato and more negative for staccato while teaching articulation. When teaching dynamics, key velocity should be higher for forte and lower for piano. Given the findings we observed in Experiment 1, we also predicted that participants would make a larger key velocity contrast between forte and piano at transition points, and that they might play more slowly when teaching, especially when teaching articulation.

2.3.1 Methods

Participants. We recruited 21 piano experts who already had a degree (above bachelor or equivalent) in piano performance/teaching or were studying advanced piano performance at a music school at the time of recruitment. For data analysis, we excluded one participant due to insufficient motor skills. Twenty participants (9 female) were included in the data analysis. Most participants were right-handed (left: 2) with a mean age of 25.90 (*SD* = 4.68). They had 15.65 years of practice on average (*SD* = 5.67). 11 participants had experience in teaching the piano (*M* = 4.14 years, *SD* = 3.65). All participants gave their informed consent before the experiment started and received vouchers for their participation. The study (No. 2018-124) was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary.

Apparatus and stimuli. The same apparatus as in Experiment 1 was used. We selected Clementi's Sonatina Op.36 (No.3) in C major as a stimulus because it contains our targeted expressions (i.e., articulation and dynamics) and is relatively simple in terms of motor skills. The first 12 measures of the original piece were used and modified so that the piece had an almost equal number of data points (i.e., the number of notes, the number of intervals between notes) for each dependent variable. The modified piece consisted of a 12-measure isochronous melody notated in a 4/4 metre to be played with the right hand only. Original sheet music was used for the purpose of practice (*Fig 2.5 A*). Expressive notations were added to the original sheet music for the experiment (*Fig 2.5 B-C*). These excerpts were confirmed to be musically natural by a doctoral student in piano performance at Liszt Ferenc Academy of Music in Hungary. The fingering was also assigned and confirmed by the same doctoral student.

Procedure. We employed the same procedure as in Experiment 1 with several modifications. First, all participants were required to memorise the piece without expressive notations (i.e., *Fig 2.5 A*) prior to the experiment so that they had enough time to familiarise themselves with the piece and perform it without pitch errors while performing it with different notated expressions in the lab. Second, we modified the wording of instructions for the performing condition so that both instructions had the same focus on expressive notations because there had been less emphasis on expression in the instruction in the performing condition in Experiment 1 (see details in *Supplementary Material 1*). Third, participants could choose their preferred tempo from one of three options (100, 110 and 120 quarter beats per minute) because some evidence shows that musicians have their own preferred tempo, which affects their temporal variability (Zamm, Wang, & Palmer, 2018). The chosen tempo was again cued by a leading metronome. To make sure that participants memorised the piece and had sufficient motor skills, we asked participants to perform the piece without looking at the original sheet music (*Fig 2.5 A*). Those who could not perform the piece without pitch errors twice consecutively within five attempts were excluded from data analysis (as a result, one participant was excluded). The rest of the procedure was identical to Experiment 1.

Data analysis. Data cleaning, preprocessing and statistical analysis were almost identical to Experiment 1. For statistical analysis, only 8th notes with expressive notations were included. As a result, only one 8th note in the 4th measure without any expression was not included. IOIs were normalised by their preferred tempo because each participant chose a tempo from the three options. Given the different tempi, key-overlap ratios (KORs) were calculated by dividing KOT by the mean IOI of each performance to normalise KOT. Additionally, we included KV difference (i.e., KV difference for each interval) at transition points (i.e., forte to piano or piano to forte) as a dependent variable based on the findings of Experiment 1. Three trials were entirely excluded from data analysis because participants did not follow the sheet music closely enough. Using the same approach as in Experiment 1, pitch errors were removed manually. For onsets, 11.62% of the trials contained at least one pitch error (extra notes: 5.81%, missing notes: 5.49%, substituted notes: 0.31%). For offsets, 17.90% of the trials contained at least one pitch error (extra notes: 5.81%, missing notes: 5.49%, substituted notes: 6.59%). As a result, less than 1 % of total responses were corrected. For each dependent variable, removing outliers (i.e., responses outside 3 standard deviations from the mean) resulted in less than 5% of overall responses being removed.

Figure 2.5. (A)Original sheet music. (B)Articulation. The curved line (slur) indicates legato and the dots indicate staccato. (C)Dynamics. The symbol 'f' denotes forte and the symbol 'p' denotes piano. For data analysis, only the 8th notes were included.

2.3.2 Results

As Experiment 1, we first report results for performance of the piece with the notated expression of articulation (*Fig 2.5 B*), followed by performance of the piece with the notated expression of dynamics (*Fig 2.5 C*).

Articulation

Inter-onset Intervals (IOIs). A paired-sample *t*-test showed that participants played more slowly in the teaching condition [*M* = 0.97, *SD* = 0.049] than in the performing condition [*M* = 0.95, *SD* = 0.040] while playing the piece with the notated articulation ($t(19) = 2.47$, $p = 0.023$, Cohen's $d = 0.55$, two-tailed, *Fig 2.6*). A Wilcoxon Signed-rank test also confirmed that there was a significant difference between the two conditions ($p = 0.033$, $r = 0.48$, two-tailed).

Key-Overlap Ratios (KORs). A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Articulation (legato vs. staccato) showed that there was a significant main effect of Articulation ($F(1,19) = 2566$, $p < 0.001$, $\eta^2_G = 0.98$) and a significant interaction between Condition and Articulation ($F(1,19) = 8.75$, $p = 0.008$, $\eta^2_G = 0.01$, *Fig 2.7*). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment showed that for staccato KORs were more negative in the teaching condition [*M* = -0.76, *SD* = 0.06] than in the performing condition [*M* = -0.74, *SD* = 0.04] (*p* = < 0.001). For legato there was no significant difference in KORs between the teaching [*M* = 0.12, *SD* = 0.07] and performing condition [*M* = 0.12, *SD* = 0.07] (*p* = 0.64).

Key Velocity (KV). A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Articulation (legato vs. staccato) showed that there was a significant main effect of Articulation ($F(1,19) = 7.85$, $p = 0.011$, $\eta^2_G = 0.03$, *Fig 2.8*), reflecting louder sound for legato notes compared to staccato notes. Neither the main effect of Condition ($F(1,19) = 1.34$, $p = 0.26$, $\eta_G^2 = 0.003$) nor the interaction between Condition and Articulation ($F(1,19) = 0.15$, $p = 0.71$, $\eta_G^2 = 0.000$) was significant.

Dynamics

Inter-onset Intervals (IOIs). A paired-sample t-test showed no significant difference between the teaching condition [$M = 0.96$, $SD = 0.048$] and the performing condition [$M = 0.96$, $SD = 0.043$] while playing the piece with the notated dynamics ($t(19) = 0.21$, $p = 0.84$, Cohen's $d = 0.05$, two-tailed, *Fig 2.6*). A Wilcoxon Signed-rank test also confirmed that there was no significant difference between the two conditions ($p = 0.96$, $r = 0.02$, two-tailed).

Key Velocity (KV). A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Dynamics (forte vs. piano) showed that there was a significant main effect of Dynamics ($F(1,19) = 131$, $p < 0.001$, $\eta_G^2 = 0.72$) and a significant interaction between Condition and Dynamics ($F(1,19) = 6.11$, $p = 0.023$, $\eta_G^2 = 0.018$, *Fig 2.8*). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment revealed that KV was higher in the teaching condition [$M = 83.20$, $SD = 8.84$] than in the performing condition [$M = 81.49$, $SD = 8.12$] ($p = 0.020$) when instructed to play forte. Also, KV was lower in the teaching condition [$M = 60.54$, $SD = 3.76$] than in the performing condition [$M = 62.43$, $SD = 4.83$] ($p = 0.046$) when instructed to play piano.

KV Transition Points. A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Transition Type (FtoP vs. PtoF) showed that there was a significant main effect of Transition Type ($F(1,19) = 126$, $p < 0.001$, $\eta_G^2 = 0.83$) and a significant interaction between Condition and Transition Type ($F(1,19) = 9.13$, $p = 0.007$, η_G^2

= 0.059, *Fig 2.9*). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment showed that there was a larger KV difference when changing from forte to piano (p = 0.028) and from piano to forte (p = 0.004) in the teaching condition [FtoP: M = -17.28, SD = 8.17, PtoF: M = 20.72, SD = 8.87] than in the performing condition [FtoP: M = -14.08, SD = 7.62, PtoF: M = 16.08, SD = 7.32].

Key-Overlap Ratios (KORs). A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Dynamics (forte vs. piano) showed that there was a significant main effect of Dynamics ($F(1,19) = 629$, $p < 0.001$, $\eta_G^2 = 0.94$), reflecting more key-overlap for forte notes compared to piano notes. Neither the main effect of Condition ($F(1,19) = 0.084$, $p = 0.77$, $\eta_G^2 = 0.000$) nor the interaction between Condition and Dynamics ($F(1,19) = 0.45$, $p = 0.51$, $\eta_G^2 = 0.001$, *Fig 2.7*) was significant.

2.3.3 Discussion

The results from Experiment 2 replicate our earlier findings and provide further evidence that skilled pianists modulated their performance in specific ways for teaching purposes. As in Experiment 1, we observed slower performance only while teaching the notated articulation. Again, KORs showed that when teaching articulation, participants exaggerated staccato while there was no significant difference in legato between the two conditions. While teaching dynamics, KV showed that participants exaggerated forte and piano. This pattern of exaggeration was more pronounced than in Experiment 1, where post-hoc comparions did not find significant differences between the two conditions for each subcomponent. Furthermore, we again found that when teaching dynamics, at transition points participants produced a larger contrast in dynamics between forte and piano, bidirectionally. Importantly, participants did not modulate performance aspects that were irrelevant for the techniques to be taught (i.e., no modulation of dynamics when teaching articulation and vice versa). Overall, Experiment 2 demonstrated systematic and fine-grained pedagogical performance modulations for a naturalistic piece of music.

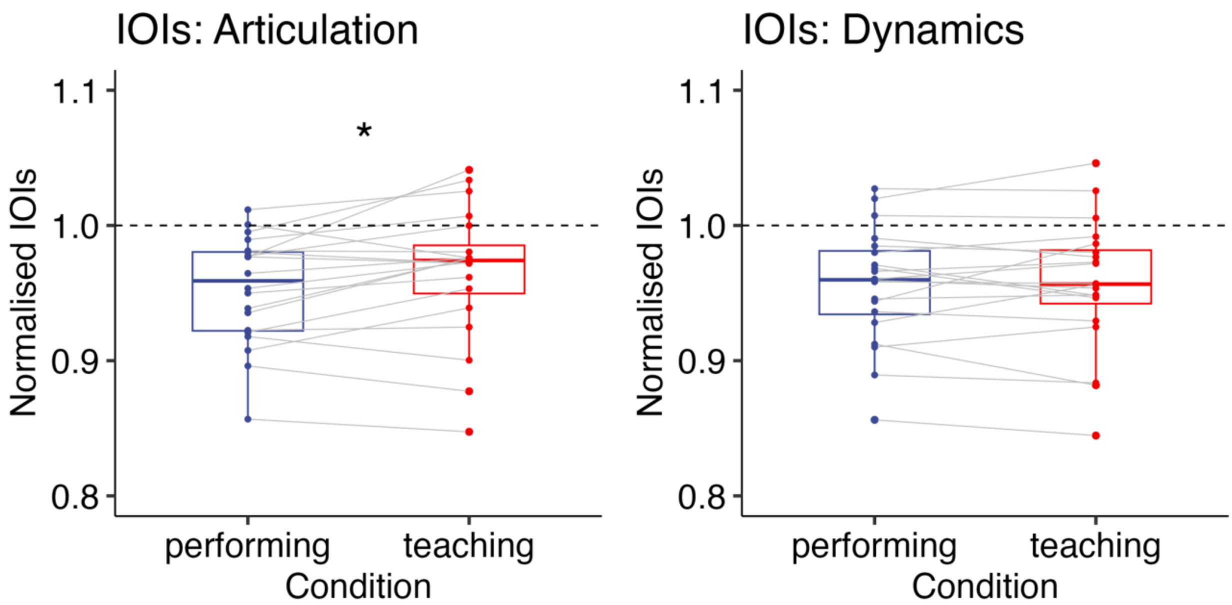

*Figure 2.6. Experiment 2: Normalised IOIs when playing the piece with either articulation (left) or dynamics (right). A dashed line represents the tempo given by a metronome. Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

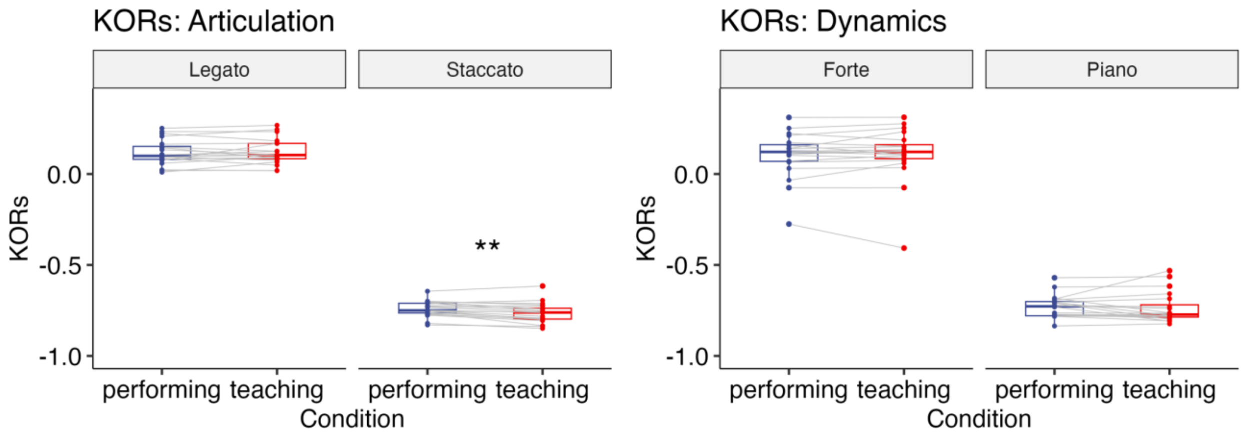

*Figure 2.7. Experiment 2: KORs when playing the piece with either articulation (left) or dynamics (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

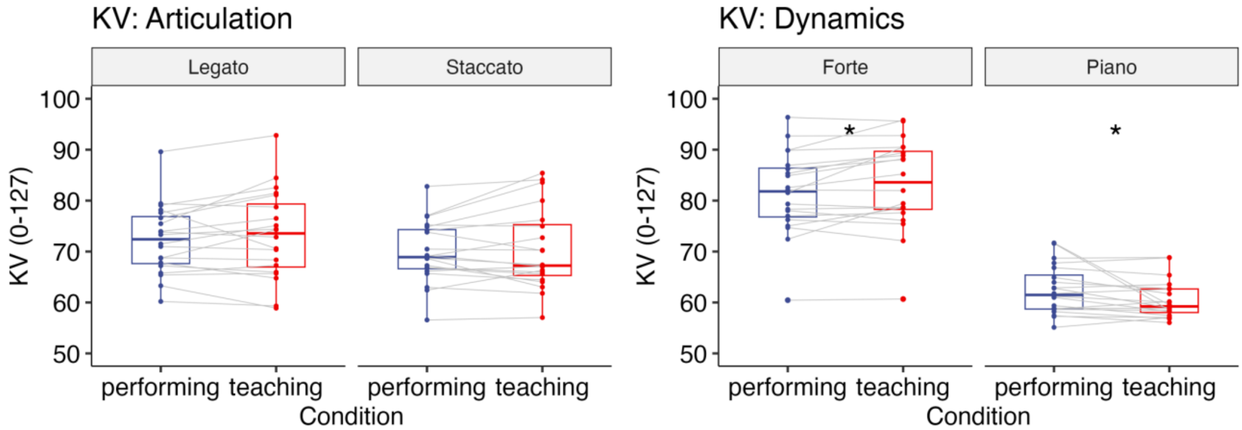

*Figure 2.8. Experiment 2: KV (0-127) when playing the piece with either articulation (left) or dynamics (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

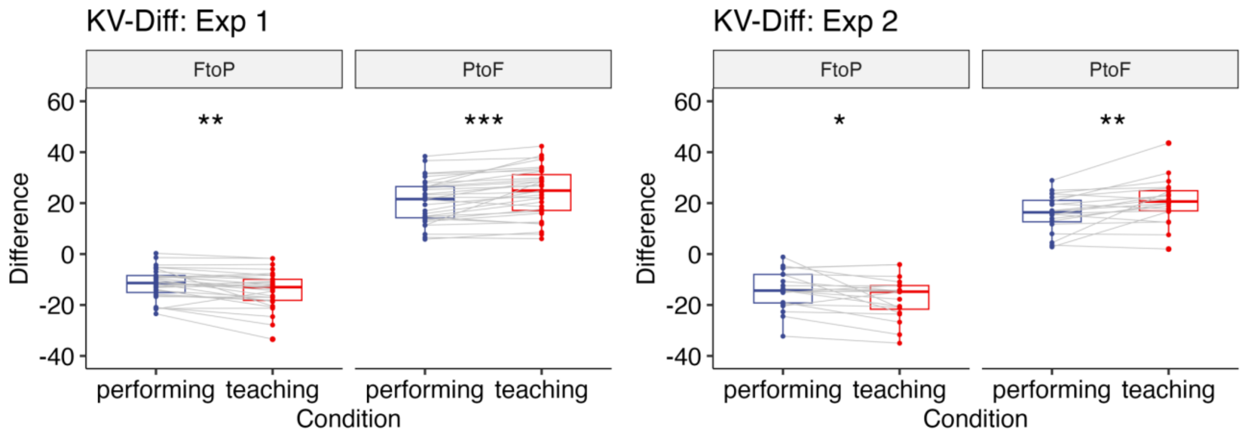

*Figure 2.9. Experiment 1, 2: KV Difference at transition points when playing the piece with dynamics (left: Experiment 1, right: Experiment 2). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

2.4 General Discussion

The aim of this study was to investigate how expert pianists modulate their performance to teach techniques necessary to implement musical expressions. Overall, the findings of the two experiments showed that pianists performed one and the same piece differently depending on whether they played with the intention to teach techniques or with the intention to perform the piece for an audience. When playing with the intention to teach, pianists selectively highlighted relevant aspects of artistic expressions. Across two different pieces, participants exaggerated staccato when playing with the intention to teach expressions concerning articulation. The lack of a difference for legato when teaching articulation might stem from a ceiling effect because there might not be enough room for exaggeration without significantly dropping the tempo. When playing with the intention to teach expressions concerning dynamics, participants made larger dynamics changes between forte and piano (in both directions). This constitutes an effective way to highlight the technique as loudness is determined relatively. Taken together, our findings demonstrate that expert pianists systematically modulated their sound depending on the specific technique they were trying to convey.

Although participants tended to play slower in the teaching condition than in the performing condition in general, we found a significant difference only for articulation, not for dynamics. As mentioned in the discussion of Experiment 1, it could be that the metronome beats given prior to each performance discouraged participants from deviating from the prescribed tempo. Another possibility is that expression is best taught when leaving the tempo unchanged when it is irrelevant. This would imply that general pedagogical modulations like slowing down might be less useful in the context of teaching techniques where timing itself can be used to add an expression to a performance.

One limitation of the study is that although we instructed participants to imagine a situation in which they were teaching musical expression to students, there was no feedback from actual students. It can be expected that teachers would also make pedagogical modulations in some domains other than acoustic properties, such as gestures if they are physically present in front of students. Moreover, given that teachers modulate their demonstration throughout ongoing interactions with learners (Okazaki, Muraoka, & Osu, 2019), future studies are needed to investigate how experts dynamically adapt their performance to their students' skill levels and demonstrated abilities. Also, some of the participants in the current study did not have teaching experience at all and even those who had teaching experience had taught for only a few years. It would be important to investigate if our findings would be observed in participants with extensive experience teaching the piano. We summarised descriptive statistics depending on participants' teaching experience (see *Supplementary Material 2*).

Another limitation of the study is that it is not certain how the results might generalise to performances of more complex pieces and pieces with fewer notations. In more complex pieces of music, multiple expressive notations could be assigned to one note or phrase. Also, it is rare that almost all the notes are assigned to one expressive notation as in our stimuli. Future studies are needed to determine whether pianists would exaggerate specific parts of a piece or prioritise one aspect over the others given multiple concurrent expressive notations and more possibilities for giving the music different interpretations.

We quantified participants' sound modulations in terms of tempo, articulation and dynamics so as to examine if and how participants modulate their sound. However, it is possible that participants show pedagogical modulations in other aspects such as temporal variability and temporal grouping (see *Supplementary Materials 3* and *4*). For example, participants tended to slow down at transition points regardless of the techniques they were teaching although overall, we found that participants played slower only when teaching the piece with the notated articulation. Future research is needed to investigate how musicians make particular modulations in relation to musical structures.

Musicians and music teachers tend to consider expressive skills as performers' most important skills (Lindström, Juslin, Bresin, & Williamon, 2003). In the present study, experts focused on the teaching of specific expressions that were notated. However, expression also has other facets (Juslin, 2003) that are not only piece-related (e.g., notated expressions) but performer- or context-related (e.g., performer's interpretation, listening contexts such as recordings or concerts). The ultimate goal of expressive performance is considered to lie in developing performers' own interpretations of music and convey affects and emotions by using expressive tools such as articulation, dynamics, tempo and timing (Meissner, 2021). The current research opens up the possibility of studying the teaching of expressive performance by demonstrating that pianists could signal pedagogical intentions without relying on verbal instructions, which are heavily used in actual teaching settings (Colprit, 2000; Woody, 2000). Given that students often learn by listening to recordings (Volioti & Williamon, 2017), it is important to investigate how performance itself could affect students' learning processes. It would also be interesting to examine how teachers' playing (i.e., demonstrations) relates to the verbal instructions they give.

The present findings extend earlier research on teaching-related action modulations. First, our study sheds light on the teaching of technical skills that are an integral part of skill acquisition in artistic domains. We showed that compared to an expressive performance baseline, experts made specific modulations to teach particular techniques. Second, the specificity of the observed modulations supports the idea that teaching comprises more than generic modulations like slowing down or overall exaggeration that may draw learners' attention. Rather, expert demonstrators seem to follow principles of relevance in communication (Sperber & Wilson, 1995), highlighting only those aspects they are intending to demonstrate. How learners benefit from the perceptual and motor cues that come with specific exaggerations, and whether understanding the teacher's intentions explicitly adds to the learning success are important questions for future research.

Data availability statement

All data is available at https://osf.io/8nbjh/.

2.5 Supplementary Material

1: Instructions

Experiment 1

In the teaching condition, participants were shown the following instruction on a computer monitor (Italic sentences were highlighted in yellow colour on a black background): "Now, play what you practised as if you were teaching it to students. Students already know how to produce the sequence of the tones and now are trying to *learn how to perform the piece expressively* by listening to your performance. *Do your best as a teacher to produce the piece according to the notation that you just practised.*"

In the performing condition, participants were shown the following instruction on a computer monitor: "Now, play what you practised as if you were performing it to an audience. *Do your best as a performer to produce the piece according to the notation that you just practised.*"

Experiment 2

In the teaching condition of Experiment 2, we gave participants the exact same instruction as in Experiment 1. In the performing condition, participants were given the following instruction on a computer monitor: "Now, play what you practised as if you were performing it to an audience. *Perform the piece expressively with your interpretation. Do your best as a performer to produce the piece according to the notation that you just practised.*"

2: Descriptive statistics depending on participants' teaching experience

We summarised descriptive statistics for those with and without teaching experience in piano. Due to the lack of statistical power, we did not perform any statistical tests.

Experiment 1

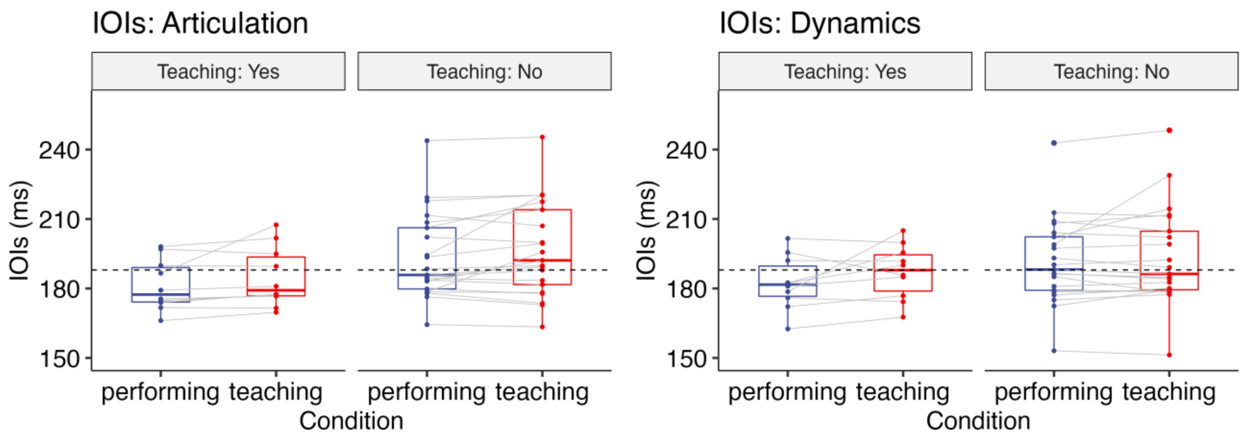

Figure S2.1. Experiment 1: Comparison between those with and without piano teaching experience in terms of IOIs (ms) when playing the piece with either articulation (left) or dynamics (right). A dashed line represents the tempo given by a metronome. Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

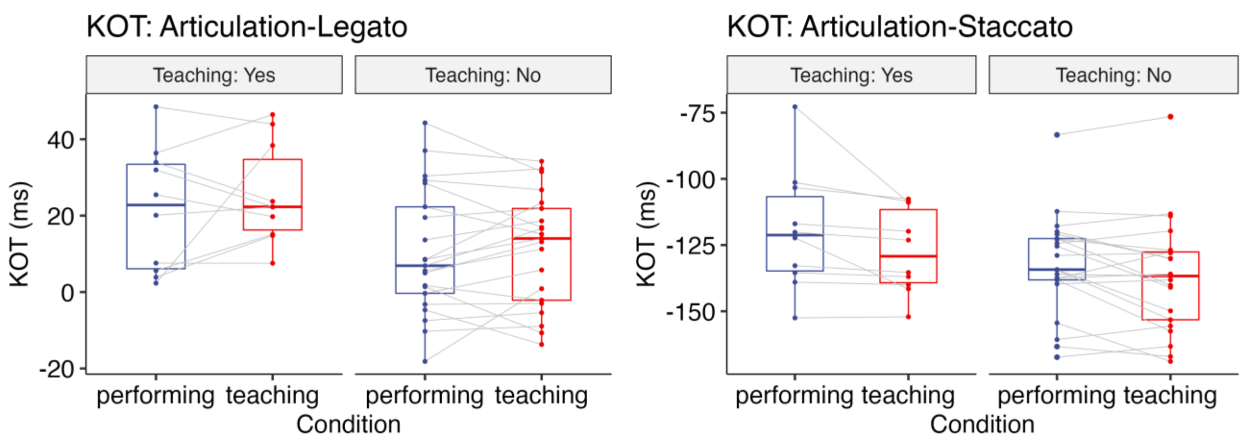

Figure S2.2. Experiment 1: Comparison between those with and without piano teaching experience in terms of KOT (ms) when playing the piece with articulation (left: legato, right: staccato). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

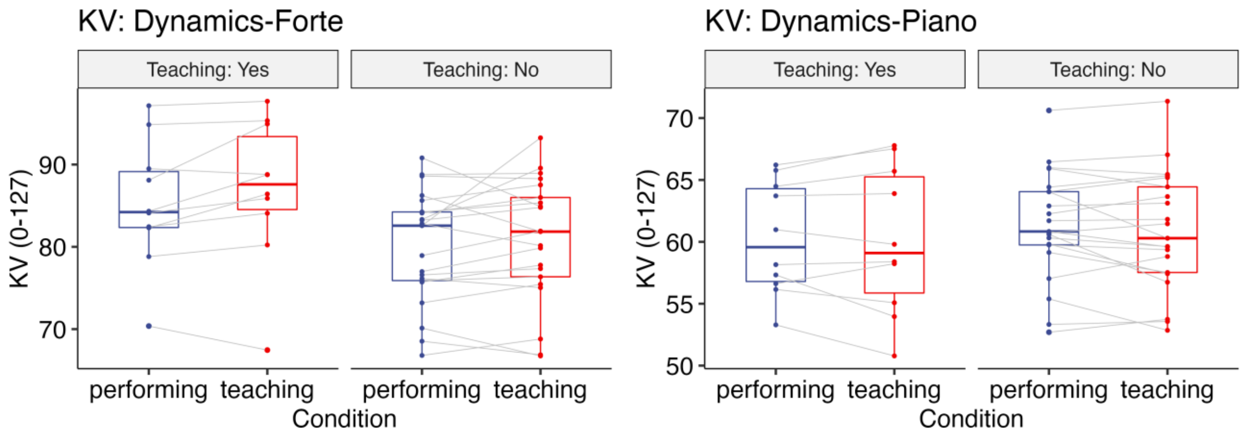

Figure S2.3. Experiment 1: Comparison between those with and without piano teaching experience in terms of KV (0-127) when playing the piece with dynamics (left: forte, right: piano). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

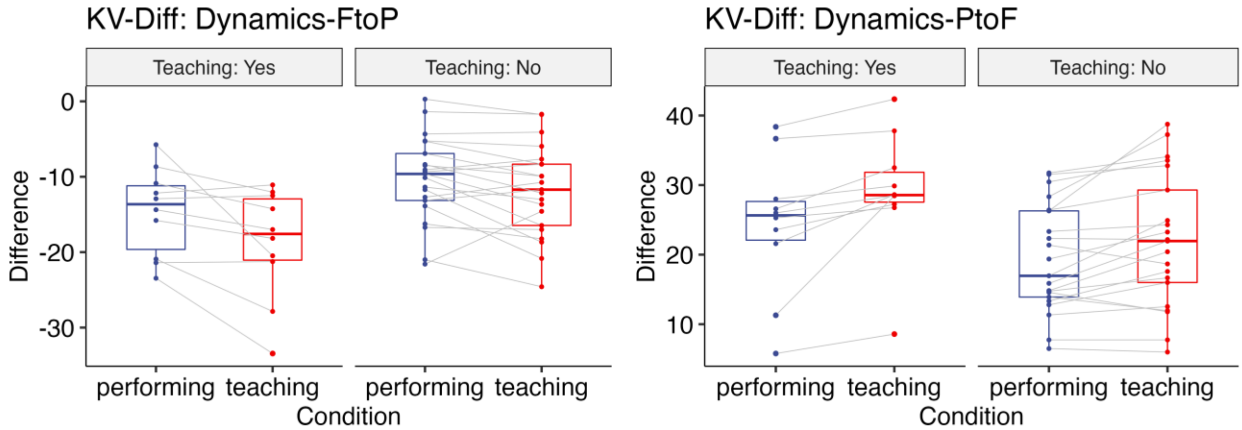

Figure S2.4. Experiment 1: Comparison between those with and without piano teaching experience in terms of KV Difference at transition points when playing the piece with dynamics (left: forte to piano, right: piano to forte). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

Table S2.1:

Experiment 1: Comparison between those with and without piano teaching experience in terms of IOIs (ms).

Condition	Technique	TeachingExp	N	Mean	SD	SEM
Performing	Articulation	Yes	10	181.27	10.99	3.47
Teaching	Articulation	Yes	10	184.72	13.02	4.12
Performing	Articulation	No	21	193.49	18.75	4.09
Teaching	Articulation	No	21	196.67	20.52	4.48
Performing	Dynamics	Yes	10	182.43	11.52	3.64
Teaching	Dynamics	Yes	10	187.15	11.73	3.71
Performing	Dynamics	No	21	191.48	18.56	4.05
Teaching	Dynamics	No	21	193.50	21.22	4.63

Note. N, SD and SEM represent sample size, standard deviation and standard error of the mean, respectively. Technique indicates performance data when participants playing the piece with either articulation (*Fig 2.1 B*) or dynamics (*Fig 2.1 C*). TeachingExp indicates if participants had experience in teaching the piano.

Table S2.2:

Experiment 1: Comparison between those with and without piano teaching experience in terms of KOT (ms).

Condition	Technique	Subcomponent	TeachingExp	N	Mean	SD	SEM
Performing	Articulation	Legato	Yes	10	21.57	16.18	5.12
Teaching	Articulation	Legato	Yes	10	25.43	13.11	4.15
Performing	Articulation	Staccato	Yes	10	-119.66	22.84	7.22
Teaching	Articulation	Staccato	Yes	10	-127.39	15.95	5.04
Performing	Articulation	Legato	No	21	10.40	16.57	3.62
Teaching	Articulation	Legato	No	21	11.36	14.71	3.21
Performing	Articulation	Staccato	No	21	-132.48	19.03	4.15
Teaching	Articulation	Staccato	No	21	-136.84	21.52	4.70

Note. N, SD and SEM represent sample size, standard deviation and standard error of the mean, respectively. Technique indicates performance data when participants play the piece with either articulation (*Fig 2.1 B*) or dynamics (*Fig 2.1 C*). TeachingExp indicates if participants had experience in teaching the piano.

Table S2.3:

Experiment 1: Comparison between those with and without piano teaching experience in terms of KV (0-127).

Condition	Technique	Subcomponent	TeachingExp	N	Mean	SD	SEM
Performing	Dynamics	Forte	Yes	10	85.20	7.76	2.45
Teaching	Dynamics	Forte	Yes	10	86.96	8.76	2.77
Performing	Dynamics	Piano	Yes	10	60.27	4.57	1.44
Teaching	Dynamics	Piano	Yes	10	60.11	5.93	1.87
Performing	Dynamics	Forte	No	21	79.97	6.82	1.49
Teaching	Dynamics	Forte	No	21	80.80	7.49	1.63
Performing	Dynamics	Piano	No	21	61.33	4.40	0.96
Teaching	Dynamics	Piano	No	21	60.88	4.76	1.04

Note. N, SD and SEM represent sample size, standard deviation and standard error of the mean, respectively. Technique indicates performance data when participants play the piece with either articulation (*Fig 2.1 B*) or dynamics (*Fig 2.1 C*). TeachingExp indicates if participants had experience in teaching the piano.

Table S2.4:

Experiment 1: Comparison between those with and without piano teaching experience in terms of KV Difference at transition points.

Condition	Technique	Subcomponent	TeachingExp	N	Mean	SD	SEM
Performing	Dynamics	FtoP	Yes	10	-14.63	5.79	1.83
Teaching	Dynamics	FtoP	Yes	10	-18.81	7.26	2.30
Performing	Dynamics	PtoF	Yes	10	24.32	9.96	3.15
Teaching	Dynamics	PtoF	Yes	10	29.07	8.78	2.78
Performing	Dynamics	FtoP	No	21	-10.30	5.69	1.24
Teaching	Dynamics	FtoP	No	21	-11.87	6.09	1.33
Performing	Dynamics	PtoF	No	21	19.23	7.71	1.68
Teaching	Dynamics	PtoF	No	21	21.98	9.56	2.09

Note. N, SD and SEM represent sample size, standard deviation and standard error of the mean, respectively. Technique indicates performance data when participants play the piece with either articulation (*Fig 2.1 B*) or dynamics (*Fig 2.1 C*). TeachingExp indicates if participants had experience in teaching the piano.

Experiment 2

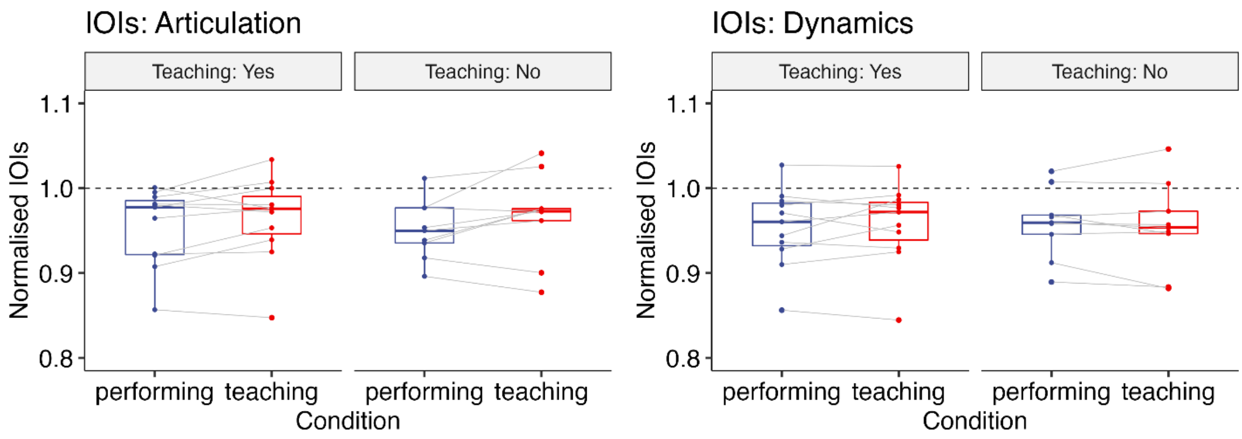

Figure S2.5. Experiment 2: Comparison between those with and without piano teaching experience in terms of normalised IOIs when playing the piece with either articulation (left) or dynamics (right). A dashed line represents the tempo given by a metronome. Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

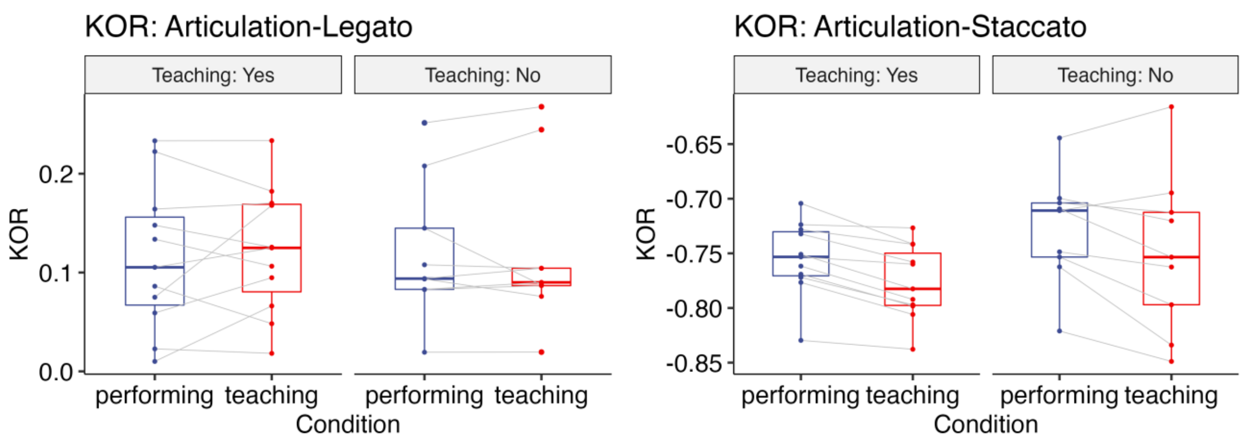

Figure S2.6. Experiment 2: Comparison between those with and without piano teaching experience in terms of KOR when playing the piece with articulation (left: legato, right: staccato). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

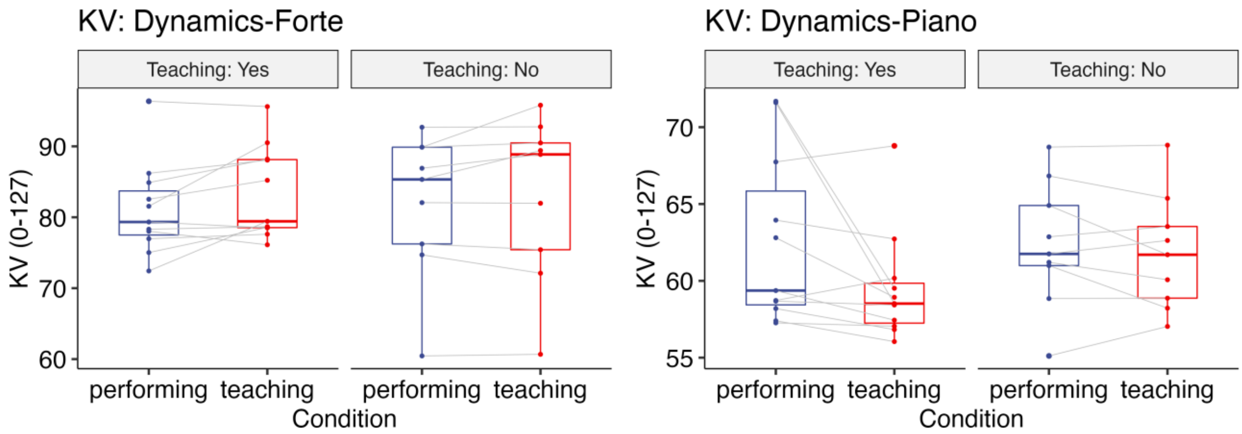

Figure S2.7. Experiment 2: Comparison between those with and without piano teaching experience in terms of KV (0-127) when playing the piece with dynamics (left: forte, right: piano). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

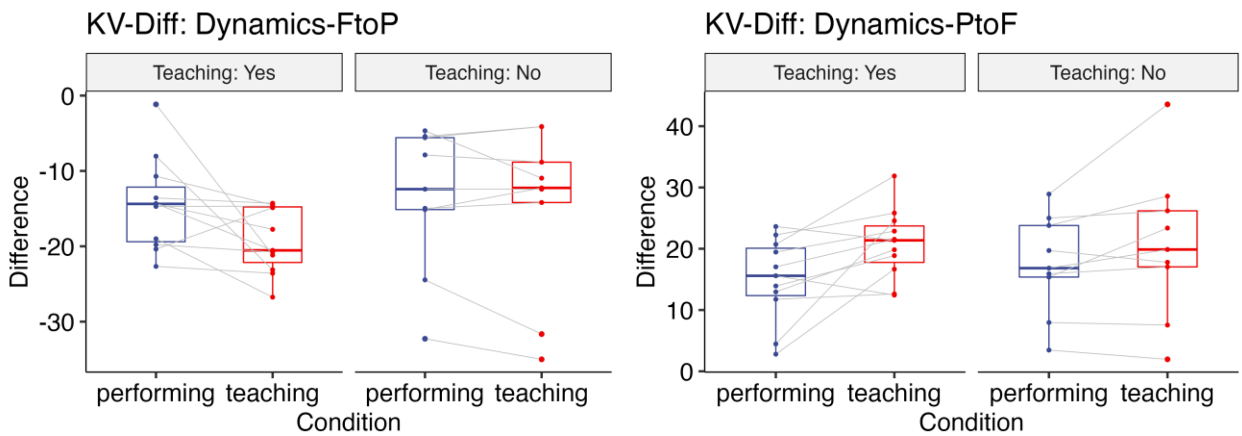

Figure S2.8. Experiment 2: Comparison between those with and without piano teaching experience in terms of KV Difference at transition points when playing the piece with dynamics (left: forte to piano, right: piano to forte). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

Table S2.5:

Experiment 2: Comparison between those with and without piano teaching experience in terms of normalised IOIs.

Condition	Technique	TeachingExp	N	Mean	SD	SEM
Performing	Articulation	Yes	11	0.954	0.046	0.014
Teaching	Articulation	Yes	11	0.964	0.049	0.015
Performing	Articulation	No	9	0.951	0.035	0.012
Teaching	Articulation	No	9	0.967	0.052	0.017
Performing	Dynamics	Yes	11	0.953	0.046	0.014
Teaching	Dynamics	Yes	11	0.958	0.047	0.014
Performing	Dynamics	No	9	0.958	0.041	0.014
Teaching	Dynamics	No	9	0.955	0.052	0.017

Note. N, SD and SEM represent sample size, standard deviation and standard error of the mean, respectively. Technique indicates performance data when participants play the piece with either articulation (*Fig 2.5 B*) or dynamics (*Fig 2.5 C*). TeachingExp indicates if participants had experience in teaching the piano.

Table S2.6:

Experiment 2: Comparison between those with and without piano teaching experience in terms of KORs.

Condition	Technique	Subcomponent	TeachingExp	N	Mean	SD	SEM
Performing	Articulation	Legato	Yes	11	0.115	0.074	0.022
Teaching	Articulation	Legato	Yes	11	0.122	0.064	0.019
Performing	Articulation	Staccato	Yes	11	-0.755	0.034	0.010
Teaching	Articulation	Staccato	Yes	11	-0.776	0.034	0.010
Performing	Articulation	Legato	No	9	0.120	0.071	0.024
Teaching	Articulation	Legato	No	9	0.120	0.081	0.027
Performing	Articulation	Staccato	No	9	-0.728	0.050	0.017
Teaching	Articulation	Staccato	No	9	-0.749	0.073	0.024

Note. N, SD and SEM represent sample size, standard deviation and standard error of the mean, respectively. Technique indicates performance data when participants play the piece with either articulation (*Fig 2.5 B*) or dynamics (*Fig 2.5 C*). TeachingExp indicates if participants had experience in teaching the piano.

Table S2.7:

Experiment 2: Comparison between those with and without piano teaching experience in terms of KV (0-127).

Condition	Technique	Subcomponent	TeachingExp	N	Mean	SD	SEM
Performing	Dynamics	Forte	Yes	11	81.06	6.52	1.97
Teaching	Dynamics	Forte	Yes	11	83.32	6.48	1.95
Performing	Dynamics	Piano	Yes	11	62.49	5.55	1.67
Teaching	Dynamics	Piano	Yes	11	59.50	3.59	1.08
Performing	Dynamics	Forte	No	9	82.02	10.15	3.38
Teaching	Dynamics	Forte	No	9	83.06	11.54	3.85
Performing	Dynamics	Piano	No	9	62.36	4.11	1.37
Teaching	Dynamics	Piano	No	9	61.81	3.76	1.25

Note. N, SD and SEM represent sample size, standard deviation and standard error of the mean, respectively. Technique indicates performance data when participants play the piece with either articulation (*Fig 2.5 B*) or dynamics (*Fig 2.5 C*). TeachingExp indicates if participants had experience in teaching the piano.

Table S2.8:

Experiment 2: Comparison between those with and without piano teaching experience in terms of KV Difference at transition points.

Condition	Technique	Subcomponent	TeachingExp	N	Mean	SD	SEM
Performing	Dynamics	FtoP	Yes	11	-14.43	6.20	1.87
Teaching	Dynamics	FtoP	Yes	11	-19.27	4.33	1.30
Performing	Dynamics	PtoF	Yes	11	14.96	6.78	2.04
Teaching	Dynamics	PtoF	Yes	11	20.77	5.71	1.72
Performing	Dynamics	FtoP	No	9	-13.64	9.46	3.15
Teaching	Dynamics	FtoP	No	9	-14.84	11.08	3.69
Performing	Dynamics	PtoF	No	9	17.44	8.11	2.70
Teaching	Dynamics	PtoF	No	9	20.66	12.09	4.03

Note. N, SD and SEM represent sample size, standard deviation and standard error of the mean, respectively. Technique indicates performance data when participants play the piece with either articulation (*Fig 2.5 B*) or dynamics (*Fig 2.5 C*). TeachingExp indicates if participants had experience in teaching the piano.

3: Temporal Variability

The tempo variability of performance was assessed with the coefficient of variation (CV) of the IOIs to investigate the tempo stability of performance between the teaching and performing condition. CVs were calculated by dividing the standard deviation by the mean of IOIs.

Results

Experiment 1

Articulation. To compare the mean CVs between the teaching and performing condition, we conducted a Wilcoxon Signed-rank test, instead of a paired *t*-test, because a Shapiro-Wilk test showed that the distribution of the mean difference was significantly different from the normal distribution (p = 0.007). The Wilcoxon Signed-rank test showed that the tempo variability of performance did not differ significantly between the teaching [*Mdn* = 0.064, *IQR* = 0.024] and performing condition [*Mdn* = 0.066, *IQR* = 0.016] while playing the piece with the notated articulation (p = 0.06, r = 0.34, two-tailed, *Fig S2.9*).

Dynamics. To compare the mean CVs between the teaching and performing condition, we conducted a Wilcoxon Signed-rank test, instead of a paired *t*-test, because a Shapiro-Wilk test showed that the distribution of the mean difference was significantly different from the normal distribution (p = 0.024). The Wilcoxon Signed-rank test showed that the tempo variability of performance did not differ significantly between the teaching [*Mdn* = 0.070, *IQR* = 0.013] and performing condition [*Mdn* = 0.074, *IQR* = 0.026] while playing the piece with the notated dynamics (p = 0.22, r = 0.22, two-tailed, *Fig S2.9*).

Experiment 2

Articulation. To compare the mean CVs between the teaching and performing condition, we conducted a Wilcoxon Signed-rank test, instead of a paired *t*-test, because a Shapiro-Wilk test showed that the distribution of the mean difference was significantly different from the normal distribution (p = 0.017). The Wilcoxon Signed-rank test showed that the tempo variability of performance did not differ significantly between the teaching [*Mdn* = 0.044, *IQR* = 0.012] and performing condition [*Mdn* = 0.047, *IQR* = 0.014] while playing the piece with the notated articulation (p = 0.07, r = 0.41, two-tailed, *Fig S2.9*).

Dynamics. A paired-sample *t*-test showed that the tempo variability of performance did not differ significantly between the teaching [M = 0.048, *SD* = 0.011] and the performing condition [M = 0.048, *SD* = 0.011] while playing the piece with the notated dynamics ($t(19)$ = 1.83, p = 0.08, Cohen's d = 0.41, two-tailed, *Fig S2.9*).

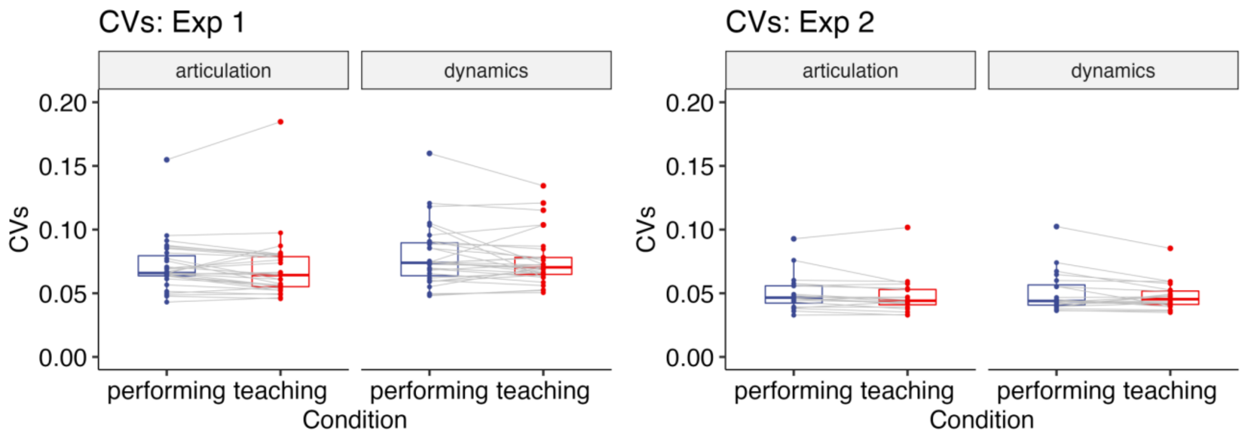

*Figure S2.9. Experiment 1, 2: CVs when playing the piece with either articulation or dynamics (left: Experiment 1, right: Experiment 2). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

4: Temporal grouping

In order to investigate if participants used temporal grouping to make boundaries between subcomponents (i.e., between legato and staccato, between forte and piano) clearer for teaching purposes, we looked at IOIs only at transition points.

Results

Experiment 1

Articulation. To compare the mean IOIs at transition points between the teaching and performing condition, we conducted a Wilcoxon Signed-rank test, instead of a paired *t*-test, because a Shapiro-Wilk test showed that the distribution of the mean difference was significantly different from the normal distribution ($p < 0.001$). The Wilcoxon Signed-rank test revealed that IOIs at transition points were larger in the teaching condition [*Mdn* = 195.75 (ms), *IQR* = 25.75] than in the performing condition [*Mdn* = 188.77 (ms), *IQR* = 23.53] while playing the piece with the notated articulation ($p = 0.02$, $r = 0.42$, two-tailed, *Fig S2.10*).

Dynamics. A paired-sample *t*-test showed that IOIs at transition points were larger in the teaching condition [$M = 200.90$ (ms), $SD = 20.96$] than in the performing condition [$M = 197.63$ (ms), $SD = 19.30$] while playing the piece with the notated dynamics ($t(30) = 2.10$, $p = 0.04$, Cohen's $d = 0.38$, two-tailed, *Fig S2.10*).

Experiment 2

Articulation. A paired-sample t-test showed that IOIs at transition points were larger in the teaching condition [M = 1.00, SD = 0.067] than in the performing condition [M = 0.98, SD = 0.049] while playing the piece with the notated articulation ($t(19) = 2.27$, $p = 0.03$, Cohen's d = 0.51, two-tailed, *Fig S2.10*).

Dynamics. To compare the mean IOIs at transition points between the teaching and performing condition, we conducted a Wilcoxon Signed-rank test, instead of a paired t-test, because a Shapiro-Wilk test showed that the distribution of the mean difference was significantly different from the normal distribution ($p = 0.004$). The Wilcoxon Signed-rank test revealed that IOIs at transition points did not differ significantly between the teaching [Mdn = 1.00, IQR = 0.06] and the performing condition [Mdn = 0.99, IQR = 0.05] while playing the piece with the notated dynamics ($p = 0.60$, $r = 0.13$, two-tailed, *Fig S2.10*).

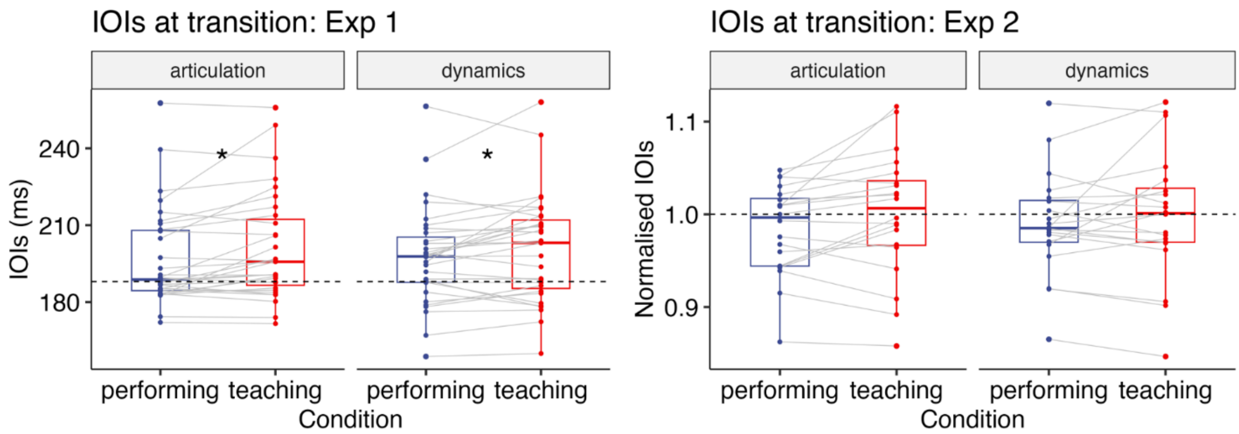

*Figure S2.10. Experiment 1, 2: IOIs at transition points when playing the piece with either articulation or dynamics (left: Experiment 1, right: Experiment 2). A dashed line represents the tempo given by a metronome. Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

5: Additional analysis for dynamics difference

In order to examine if the dynamics difference between forte and piano at transition points is larger when the difference between forte and piano in overall performance is considered, we normalised the KV difference at transition points by dividing it by the average range between forte and piano per performance (per trial).

Results

Experiment 1

A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Transition Type (FtoP vs. PtoF) showed that there was a significant main effect of Transition Type ($F(1,30) = 741$, $p < 0.001$, $\eta_G^2 = 0.88$) and a significant interaction between Condition and Transition Type ($F(1,30) = 7.87$, $p = 0.009$, $\eta_G^2 = 0.009$, *Fig S2.11*). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment showed that there was no significant difference when changing from forte to piano ($p = 0.085$) and from piano to forte ($p = 0.075$) between the teaching condition [FtoP: $M = -0.62$, $SD = 0.24$, PtoF: $M = 1.10$, $SD = 0.34$] and performing condition [FtoP: $M = -0.56$, $SD = 0.25$, PtoF: $M = 1.03$, $SD = 0.40$].

Experiment 2

A two-way repeated-measures ANOVA with the factors Condition (teaching vs. performing) and Transition Type (FtoP vs. PtoF) showed that there was a significant main effect of Transition Type ($F(1,19) = 42$, $p < 0.001$, $\eta_G^2 = 0.63$). However, there was no significant interaction between Condition and Transition Type ($F(1,19) = 1.10$, $p = 0.31$, $\eta_G^2 = 0.008$, *Fig S2.11*).

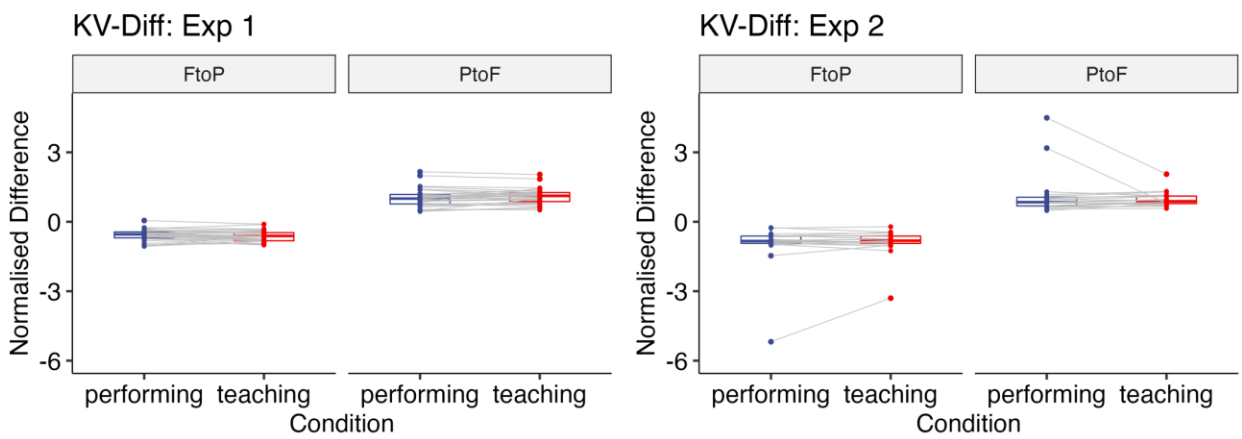

*Figure S2.11. Experiment 1, 2: Normalised KV Difference at transition points when playing the piece with dynamics (left: Experiment 1, right: Experiment 2). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box. Significance levels: * < .05, ** < .01, *** < .001*

Chapter 3.

Experts' adaptation depending on novices' skills

3.1 Introduction

Effective teaching involves experts and novices dynamically interacting with each other (Byrne & Rapaport, 2011). First, experts often demonstrate what novices are meant to acquire and then novices attempt to reproduce what the experts demonstrated. During a skill acquisition process, experts regularly monitor novices' actions and intervene by adapting their behaviour to novices' observed abilities (Mermelshtine, 2017). This interactive learning process will be iterated until novices can perform independently without the help of experts. Experts scaffold novices' learning by identifying novices' specific problems from their unsuccessful attempts and modifying their behaviour accordingly (Ugur, Nagai, Celikkanat, & Oztop, 2015; Zukow-Goldring & Arbib, 2007).

Exaggeration of movement is often used to signal a communicative intent to others during real-time interactions (Pezzulo et al., 2019), including teaching contexts. When mothers were teaching how to use a novel toy to their infants, they tended to demonstrate slowly and to produce wide and expansive movements compared to when they were demonstrating to adults (Brand et al., 2002). Such pedagogical action modulations are sensitive to the skill level of learners. For instance, evidence from parental scaffolding research suggests that caregivers scaffold infants' learning when the infants failed to produce an expected action (Wood & Middleton, 1975; Zukow-Goldring & Arbib, 2007). Fukuyama et al. (2015) performed a study, where mothers were asked to demonstrate a cup-nesting task to their infants. There were four cups of different colours and sizes. Mothers demonstrated how to stack and fit the nesting cups with the intention to teach their infants to reproduce the actions by themselves. It was found that mothers dynamically modulated their demonstrations according to the infants' object manipulation skills. Specifically, mothers exaggerated their movements when their infants failed to produce expected actions but did not produce such exaggerations when their infants lacked the motor skills to produce the required actions. Movement exaggerations for the purpose of teaching are observed not only towards infant learners but also when knowledgeable adults are interacting with adult learners (McEllin et al., 2017). For example, participants demonstrating to a learner how to play a particular melody on a xylophone increased the amplitude of their movements, compared to playing the same melody on their own. However, it has not been established whether and how experts teaching skills to adult learners adjust their demonstrations as a function of the learners' demonstrated skills.

To address this question, we investigated whether and how experts can exaggerate specific parameters of the action to be acquired depending on learners' skill levels. We chose to study expressive music performance, in which experts need to modulate multiple performance parameters such as articulation and dynamics to implement expression (Cancino-Chacón, Grachten, Goebl, & Widmer, 2018). Musicians can create their unique expression and interpretation of music even if they play exactly the same piece because when and how they change multiple performance parameters is different. By studying expressive performance, we can investigate how precisely experts can highlight specific parameters in a multidimensional space of expression (e.g., tempo, articulation, dynamics) depending on the goal of teaching (e.g., to teach articulation or dynamics) even when they are performing the same piece.

In our previous research (Tominaga, Knoblich, & Sebanz, 2022; Chapter 2), we demonstrated that expert pianists modulated their performance when they were asked to teach musical expressive techniques such as articulation and dynamics. In particular, expert pianists exaggerated the contrast between different dynamics (i.e., they made a larger contrast between forte and piano) when they had the intention to demonstrate dynamics. When they had the intention to demonstrate articulation, they exaggerated articulation by producing shorter staccato. In that study, experts were instructed to teach techniques to hypothetical learners and they did not hear novices' particular performances. This leaves open the question of whether experts would make specific adjustments depending on learners' skills in the way mothers have been shown to adapt demonstrations of actions to their infants' skills. If experts modulate their demonstrations based on learners' skills, their exaggeration of particular aspects should depend on the learners' actual performance with regard to the techniques to be acquired. For example, if learners did not implement notated articulation during their performance, experts should exaggerate the performance relevant to articulation. However, if learners implemented notated articulation correctly, experts should not exaggerate the performance relevant to articulation.

To obtain experimental control over learners' demonstrated proficiency in implementing particular techniques, we generated artificial recordings of novice piano performances instead of recruiting actual novice pianists. The recordings were generated based on the performance data from our previous study (Tominaga et al., 2022). All the recordings were of the same piece but the recordings differed in the implementation of articulation and dynamics. According to the notated expressions in the sheet music, the proper performance of the piece required specific modulations with respect to both articulation and dynamics (*Fig 3.1*). We generated four types of recordings. In one quarter of the recordings, both techniques were implemented; in one quarter, neither articulation nor dynamics was implemented; in the two remaining quarters, the piece was played with either only articulation or only dynamics implemented. This allowed us to investigate whether expert pianists in the role of teachers would make specific adjustments in response to the learners' demonstrated skills.

We hypothesised that expert pianists would exaggerate their performance only when specific techniques (i.e., either articulation or dynamics, or both) were missing in the recordings. More precisely, expert pianists would exaggerate articulation (i.e., producing longer legato and shorter staccato) when articulation was not implemented in the recordings whereas expert pianists would exaggerate dynamics (i.e., producing louder forte and softer piano, and larger contrast between forte and piano) when dynamics was not implemented in the recordings. If both articulation and dynamics were missing, we predicted that expert pianists would exaggerate both aspects.

3.2 Methods

Participants We recruited 20 participants (15 female) who already had a degree (above bachelor's or equivalent) in piano performance/teaching or were studying advanced piano performance at a music school. Most participants were right-handed (left: 2, ambidextrous: 2). The mean age of the participants was 28.25 years (*SD* = 10.95). They had 21.55 years of practice on average (*SD* = 11.59). 17 participants had teaching experience in piano (*M* = 7 years, *SD* = 6.68). All participants were recruited through an online participant platform (SONA system, https://www.sona-systems.com). The study (No. 2020/05) was approved by CEU PU's Psychological Research Ethics Board (PREBO).

Apparatus and stimuli The experiment was programmed in Max/MSP (8.1.11; https://cycling74.com/products/max) on a MacBook Pro with Mac OS X Catalina 10.15.7. A weighted Yamaha MIDI digital piano was used to record participants' performances. The pitch, onset and offset time of each note, and key velocity profiles were obtained from MIDI data using Max/MSP patchers. The laptop and piano were connected to a high-fidelity soundcard (Focusrite Scarlett 6i6) to deliver a metronome and piano sound. All auditory feedback was given to participants through headphones (Audio-Technica ATH-M50X). Sheet music was displayed on a computer monitor in front of the participants.

The recorded performances serving as stimuli were all of one piece of music that had already been used in our previous study (Tominaga et al., 2022). The piece was an excerpt from Clementi's Sonatina Op.36 (No.3) in C major. The first 12 measures of the original piece were used and modified so that the piece had an almost equal number of data points (i.e., the number of notes, the number of intervals between notes) for each dependent variable. The modified piece consisted of a 12-measure isochronous melody notated in 4/4 meter to be played with the right hand only. Two expressive notations (i.e., articulation and dynamics) were added, as shown in *Fig 3.1*. The fingering was assigned but participants were told that they did not have to follow the indicated fingering if they were not comfortable with it.

We created artificial recordings of novices to manipulate the displayed proficiency with regard to the two relevant techniques. There were four types of recordings: 1) recordings where both articulation and dynamics were implemented (both), 2) recordings where only articulation was implemented whereas dynamics was missing (art-only), 3) recordings where only dynamics was implemented whereas articulation was missing (dyn-only), and 4) recordings where neither articulation nor dynamics was implemented (none). We generated four instances for each type, therefore there were 16 stimuli (i.e., recordings) in total. How we generated the recordings is described in *Supplementary Material*.

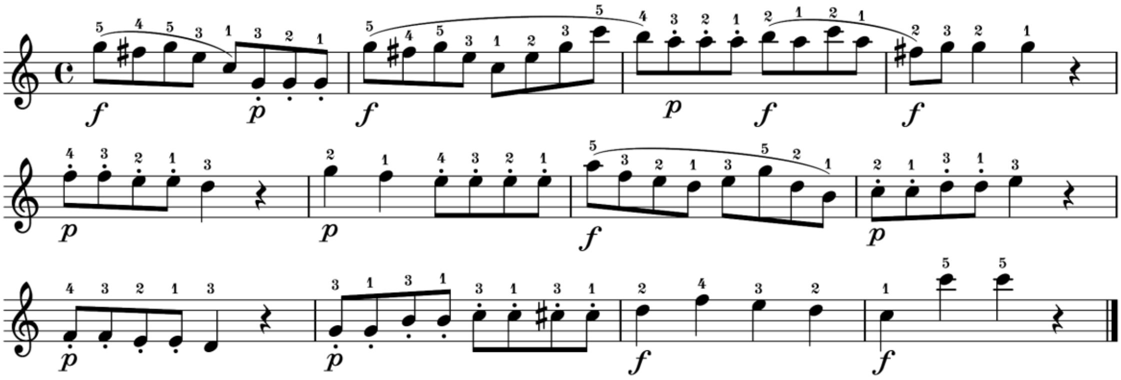

Figure 3.1. Sheet music. For articulation notation, the curved line (slur) indicates legato and the dots indicate staccato. For dynamics notation, the symbol 'f' denotes forte and the symbol 'p' denotes piano. For data analysis, only the 8th notes with expressive notations were included.

Procedure Prior to the experiment, participants were required to memorise the piece so that they could perform it with the notated expressions and without pitch errors in the experiment.

First, we recorded participants' baseline performance by asking them to perform the piece without having listened to any novices' performance. A leading metronome (100 quarter beats per minute, 8 beats) was given before participants started performing the piece. Sheet music (*Fig 3.1*) was displayed in front of the participants. They were told to perform the piece expressively with their interpretation and to do their best as a performer. This instruction was given to make sure that they paid attention to the expressive aspects of the performance but did not have the intention to teach. Each participant performed the piece twice.

After we recorded the participants' baseline performance, participants were told that they were going to listen to a number of recordings from 16 different students, who were learning musical expressive techniques. Participants were required to listen to each student's recording first, and then to perform the same piece to teach musical expressive techniques to that student. In total, there were 16 trials and participants played the piece for each student only once. The order of the recordings was randomised for each participant. A leading metronome (100 quarter beats per minute, 8 beats) was given before participants started performing the piece.

After participants completed the 16 trials, they were asked to perform the piece with the same instruction as in their baseline performances. At the end of the experiment, participants filled in a questionnaire asking about their demographic information and experience in piano performance/teaching. All participants were fully debriefed at the end of the experiment and informed that the recordings had been artificially generated.

Data analysis The dependent variables were computed from MIDI data for data analysis. Key-overlap time (KOT) is the difference between the offset time (i.e., key release time) of the current tone and the onset time of the ensuing tone and is a measure for the smoothness of musical sequences. A positive value indicates smooth legato style due to overlap between the current and ensuing tone whereas a negative value indicates sharp staccato style due to separation between the current and ensuing note. Tone intensity is assessed by key velocity (KV) and measures the loudness of a musical note. A higher value indicates forte style whereas a lower value indicates piano style. The value of KV in MIDI varies between 0 (minimum) and 127 (maximum). Also, KV difference was calculated by subtracting the KV value of the current note from that of the following note. We particularly focused on specific points where each subcomponent changed from one to the other (i.e., forte to piano or piano to forte) to measure dynamics contrast between forte and piano. Interonset intervals (IOIs) are the intervals between onsets of adjacent notes and provide a measure of tempo.

Data processing and statistical analysis were performed in R version 4.0.5. For statistical analysis, we included 8th notes with expressive notations only. Pitch errors were identified by comparing the sequence of musical notes produced by a participant with the sequence of musical notes according to the sheet music. Pitch errors included extra, missing or substituted tones and were manually removed by using the *editData* R package. For note onsets, 31.87 % of the trials contained at least one pitch error (extra notes: 5.94 %, missing notes: 24.38 %, substituted notes: 1.56 %). For note offsets, 35.31 % of the trials contained at least one pitch error (extra notes: 5.94 %, missing notes: 24.38 %, substituted notes: 5 %). We found that some participants did not precisely follow the sheet music (e.g., they held some notes longer than notated), therefore the order of offsets did not correspond to that of onsets. We considered these as errors and removed the erroneous notes even if the order of onsets was correct. As a result, less than 1 % of total responses were removed. After removing pitch errors, we removed outliers for KOT, KV, KV Difference and IOIs, defined as values more than 3 standard deviations from the mean of each dependent variable. For each dependent variable, this resulted in less than 2 % of overall responses being removed as outliers.

First, we performed a 2 x 2 repeated-measures analysis of variance (ANOVA) with the factors Articulation (present vs. absent) and Dynamics (present vs absent) for each dependent variable (i.e., KOT, KV, KV Difference, IOIs). The *aov_ez* function in the *afex* R package was used for a repeated-measures ANOVA. For post-hoc comparisons on the estimated marginal means, we used the *emmeans* R package.

In order to investigate how their performances during the experiment differed from their baseline performances (i.e., performances that participants produced for non-teaching purposes), we performed one-way repeated-measures ANOVAs. For KOT, KV and KV Difference, we performed ANOVAs separately for each subcomponent (i.e., legato, staccato, forte, piano). The results are described in *Supplementary Material*.

3.3 Results

All effects are reported as significant at $p < .05$. For KOT, KV and KV Difference, we performed two-way ANOVAs separately for each subcomponent (i.e., legato, staccato, forte, piano).

KOT

Legato. As predicted, experts produced longer legato when listening to the recordings where articulation was not implemented (*Fig 3.2*, left). Accordingly, there was a significant main effect of Articulation ($F(1, 19) = 5.59$, $p = 0.029$, $\eta_G^2 = 0.002$). There was no significant main effect of Dynamics ($F(1, 19) = 0.06$, $p = 0.81$, $\eta_G^2 = 0.000$) and no significant interaction between Articulation and Dynamics ($F(1, 19) = 1.89$, $p = 0.19$, $\eta_G^2 = 0.002$).

Staccato. In line with our prediction, experts produced shorter staccato when listening to the recordings where articulation was not implemented (*Fig 3.2*, right). Accordingly, there was a significant main effect of Articulation ($F(1, 19) = 4.88$, $p = 0.040$, $\eta_G^2 = 0.009$). There was no significant main effect of Dynamics ($F(1, 19) = 0.68$, $p = 0.42$, $\eta_G^2 = 0.000$) and no significant interaction between Articulation and Dynamics ($F(1, 19) = 3.13$, $p = 0.093$, $\eta_G^2 = 0.001$).

KV

Forte. Contrary to our predictions, experts did not change their performances in terms of notated forte depending on the type of recordings (*Fig 3.3*, left). Neither the main effect of Articulation ($F(1, 19) = 1.66$, $p = 0.21$, $\eta_G^2 = 0.00$) nor the main effect of Dynamics ($F(1, 19) = 0.33$, $p = 0.57$, $\eta_G^2 = 0.000$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 0.00$, $p = 0.96$, $\eta_G^2 = 0.000$).

Piano. Experts produced softer piano only when listening to the recordings where articulation was implemented (*Fig 3.3*, right). There was a significant main effect of Articulation ($F(1, 19) = 7.18$, $p = 0.01$, $\eta_G^2 = 0.060$). However, there was no significant main effect of Dynamics ($F(1, 19) = 1.80$, $p = 0.20$, $\eta_G^2 = 0.003$) and no significant interaction between Articulation and Dynamics ($F(1, 19) = 0.04$, $p = 0.85$, $\eta_G^2 = 0.000$).

KV Difference

Contrary to our predictions, experts did not change their performances in terms of dynamics contrast depending on the type of recordings (*Fig 3.4*).

Forte to Piano. Neither the main effect of Articulation ($F(1, 19) = 3.40$, $p = 0.081$, $\eta_G^2 = 0.012$) nor Dynamics ($F(1, 19) = 1.58$, $p = 0.22$, $\eta_G^2 = 0.002$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 0.97$, $p = 0.34$, $\eta_G^2 = 0.001$).

Piano to Forte. Neither the main effect of Articulation ($F(1, 19) = 2.25$, $p = 0.15$, $\eta_G^2 = 0.005$) nor Dynamics ($F(1, 19) = 3.76$, $p = 0.07$, $\eta_G^2 = 0.004$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 1.08$, $p = 0.31$, $\eta_G^2 = 0.001$).

IOIs

Contrary to our predictions, experts did not change the tempo depending on the type of recordings (*Fig 3.5*). Neither the main effect of Articulation (F(1, 19) = 2.79, p = 0.11, η^2_G = 0.010) nor the main effect of Dynamics (F(1, 19) = 3.27, p = 0.086, η^2_G = 0.003) nor the interaction between Articulation and Dynamics was significant (F(1, 19) = 1.48, p = 0.24, η^2_G = 0.002).

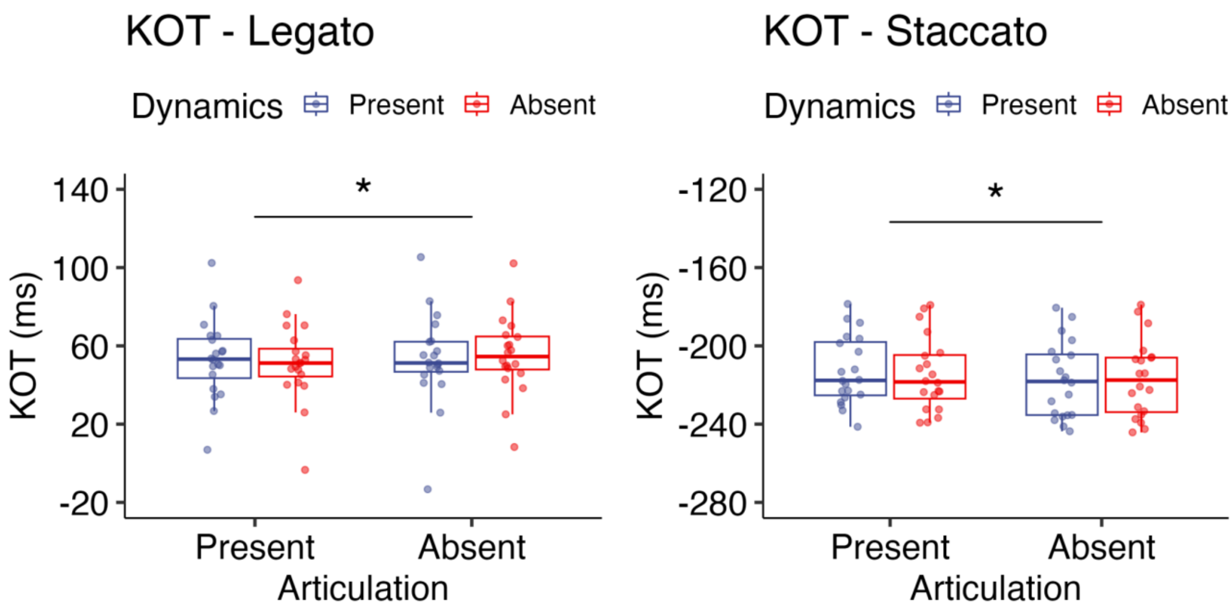

Figure 3.2. KOT(ms) for each subcomponent; legato (left) and staccato (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

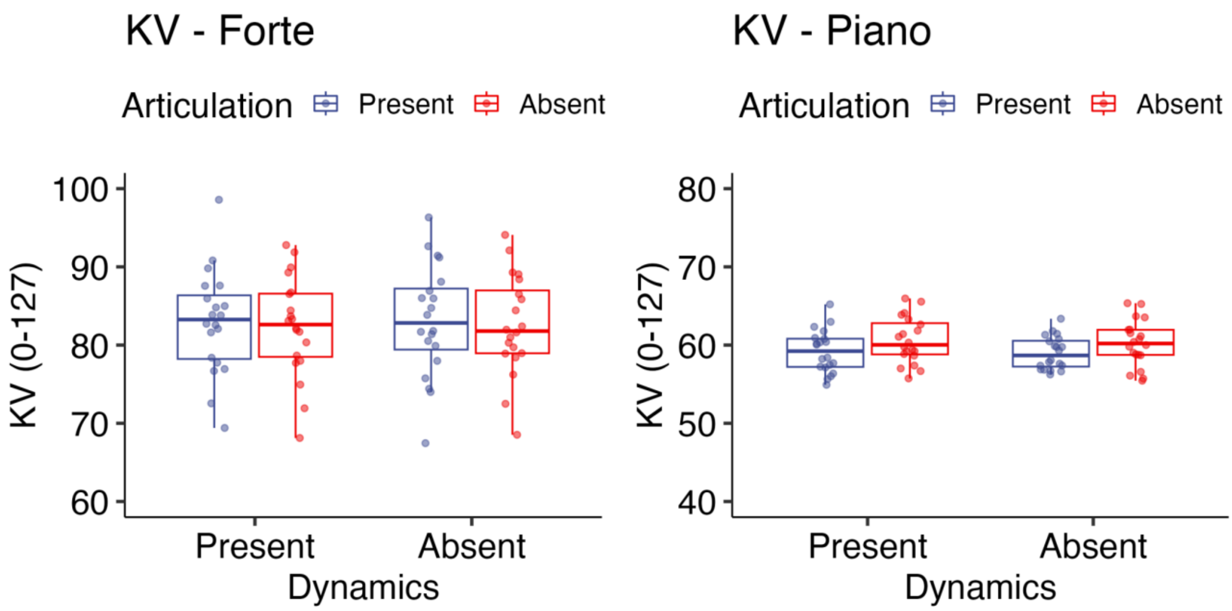

Figure 3.3. *KV (0-127) for each subcomponent; forte (left) and piano (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.*

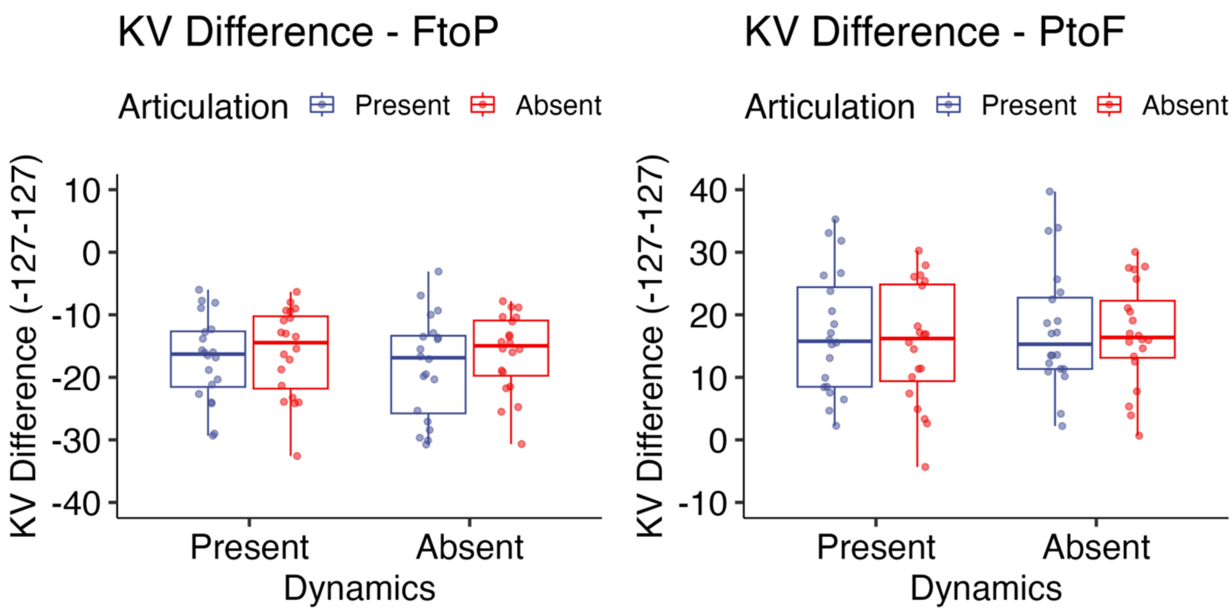

Figure 3.4. *KV Difference (-127-127) for each subcomponent; forte to piano (left) and piano to forte (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.*

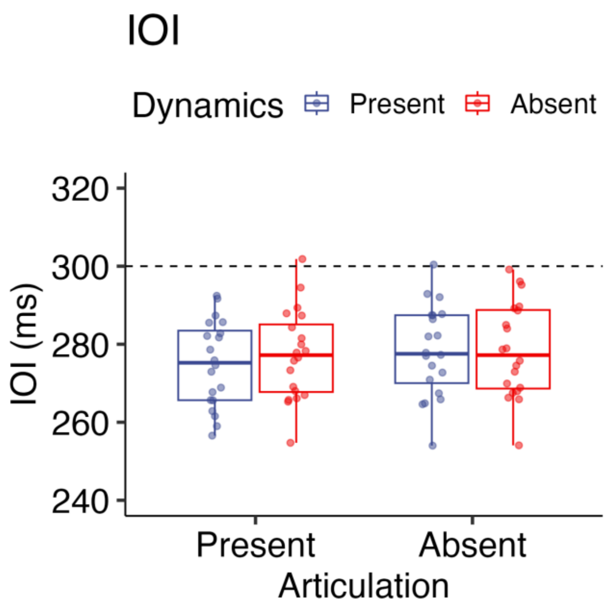

Figure 3.5. IOIs (ms). A dashed line represents the tempo given by a metronome. Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

3.4 Discussion

The present study investigated whether and how expert pianists adapt their performance depending on the displayed skills of novices whom they are intending to teach. We created artificial recordings to manipulate the implementation of two musical expressive techniques (i.e., articulation and dynamics) notated on the sheet music. The recordings where the two techniques were implemented were supposed to resemble the performances of students with good skills. The recordings where neither of the two techniques was implemented were supposed to resemble the performances of students with lower skills who had not yet learnt how to implement articulation and dynamics. The recordings where either of the two techniques was implemented but the other was missing were created to examine whether expert pianists could recognise a particular problem that students have (e.g., only articulation was missing) and adjust their performance accordingly (e.g., expert pianists should modulate their performance relevant to articulation, but not dynamics).

Expert pianists exaggerated legato and staccato when articulation was not implemented in the recordings. This is in line with our predictions that experts would exaggerate only the relevant aspects of the performance if a particular technique was missing in the recordings. However, we did not find significant results in terms of dynamics. Instead, we found that expert pianists produced softer piano when articulation was implemented in the recordings, indicating that participants exaggerated piano only when there was no problem with regard to articulation in the recordings. This may imply that participants prioritised the teaching of articulation over that of dynamics and did not modify two performance parameters at the same time. This is consistent with the findings in parental scaffolding where caregivers help infants acquire a complex action sequence by simplifying the sequential action into discrete sub-actions and adding complexity step by step (Zukow-Goldring & Arbib, 2007).

The reason why expert pianists seemed to prioritise the teaching of articulation might be that articulation was more important for the specific piece we selected or that the piece itself naturally invited some implementation of articulation. In our previous study, we observed that most participants implemented articulation (particularly legato) when they were asked to perform the piece even when neither articulation nor dynamics was notated on the sheet music. Therefore, it is possible that expert pianists were particularly sensitive to the lack of articulation in the recordings and were trying to highlight it. Another possibility is that dynamics modulations in the recordings were too subtle to be noticed by participants and therefore they did not modulate the dynamics aspect of their performance.

We also found that expert pianists did not modulate their tempo (IOIs) depending on the displayed skills of students. Also compared with the tempo of the baseline performance (see *Supplementary Material*), they did not change the tempo in the teaching context. These findings indicated that tempo was not employed specifically in teaching, likely because tempo itself was not relevant to either articulation or dynamics. Although slower demonstration is generally considered to be used for teaching (Brand et al., 2002; McEllin et al., 2017; Schaik et al., 2019), in music performance, it may not be an effective strategy to perform slowly as changing the tempo might give another interpretation to the music.

It is possible that it may be difficult in general for experts to flexibly adjust to different action parameters while teaching or they might have chosen to focus on the most salient parameters or the ones that they found most relevant. For example, when teaching dance, it is considered that good teachers identify and build the basic knowledge and skills first and then add complexity progressively from the basic foundation (Mainwaring & Krasnow, 2010). When teaching social dance, a teacher started with a basic step as the hustle and then moved on more complex steps progressively like the waltz and the cha cha (Lobel, 2021). Karl Leimer, a German music teacher and pianist, suggested in his book that students should perform with legato at first in finger exercises such as the playing of scales before learning other techniques (Gieseking & Leimer, 1972). In both dance and music examples, the ultimate goal of skill acquisition is that students are able to use multiple techniques flexibly while performing. However, it seems that students need to build complex skills progressively based on the basics during the learning process. Therefore, identifying and focusing on the most relevant problem for each student may be more effective to facilitate skill acquisition.

One of the limitations of the current study is that participants may not have had a clear idea about what to teach because they did not have a normative ideal performance in mind.

As some participants reported that "(students had) very different interpretations of the melody" or "all kids played rhythmically correct" in the questionnaire, it is plausible that the lack of articulation and/or dynamics might not have been perceived as errors. One solution would be to provide a model performance where both articulation and dynamics are implemented and ask participants to listen to it first so that they can detect errors by comparing each recording with the model performance.

The reason why participants considered different recording variations as interpretations rather than errors might stem from the fact that none of the recordings included any pitch errors and all were performed with a stable tempo. This was to make sure that only articulation and dynamics features of the recordings varied. However, some of the recordings might not have sounded as if unskillful students were playing. In order to create more realistic novice performances, it might be useful to add some jitter to the tempo of the recordings we created, or to ask beginner pianists to play the piece and extract the temporal features of their performances.

Experiments which investigated dynamic interactions between experts and novices have been done with the physical presence of both an expert and a novice in the same place (Fukuyama et al., 2015; Okazaki et al., 2019; Zukow-Goldring & Arbib, 2007). In the current experiment, expert pianists only had access to their imaginary students and needed to perform in the absence of actual students. Future research should investigate how expert pianists and novices dynamically interact in a situation where they can communicate in real-time and examine how the performances of expert pianists and those of novices are related to each other.

Data availability statement

All data is available at https://osf.io/becf6/.

3.5 Supplementary Material

1: Artificial recordings

Artificial recordings were generated from performance data that we obtained in our previous study (Tominaga et al., 2022). We used the data from Experiment 2. In Experiment 2, participants were asked to memorise one piece of music (*Fig 5* (A), Tominaga et al. (2022)) without any expressive notation before coming to the lab. First, we asked participants to perform the memorised piece before starting the experiment to measure their baseline performances. After measuring the baseline performances, participants were asked to perform the piece with either articulation (*Fig 5* (B), Tominaga et al. (2022)) or dynamics (*Fig 5* (C), Tominaga et al. (2022)) in two conditions (teaching vs. performing). Here, we only used the data from the performing condition to make sure that the performances we used did not have any particular modulations because of teaching intentions. Therefore, there were two types of performance data (articulation performance: performances played with notated articulation, dynamics performance: performances played with notated dynamics). We obtained 31 valid performances (i.e., performances without pitch errors) as baseline performances, 137 valid performances as articulation performances and 139 valid performances as dynamics performances.

Procedure. We aimed to generate four types of recordings (i.e., both, art-only, dyn-only, none) and each type had four instances (see *Apparatus and stimuli*) by using the performance data from Tominaga et al. (2022). First, 16 instances for each type were generated. Second, for each type, we chose the four best instances, which exhibited the characteristic of each type well.

Generation of 16 instances per each type. In order to generate artificial recordings, we needed information about 1) the pitch order (the order of each key press), 2) the onset time of each key press (related to tempo), 3) the offset time of each key press (related to articulation) and 4) the key velocity of each key press (related to dynamics). Pitch information was already determined and the same for all instances based on sheet music (*Fig 3.1*).

First, we created 16 temporal structures of recordings from the baseline performances. Three performances out of the 31 valid baseline performances were randomly selected and interonset intervals (IOIs) of the three performances were averaged for each interval. Due to the dataset size (i.e., 31 performances), some baseline performances were used twice. We made sure that the identical performance was not selected within each random selection of three performances.

Next, we were going to follow the same procedure as we determined tempi to determine the sound duration of each note (related to articulation) and the key velocity profile of each note (related to dynamics). For the sound duration, however, we found that most participants implemented articulation (particularly legato) even in their baseline performances, where they should not implement any articulation or dynamics. Therefore, we decided to use the data from only two participants who did not show such tendencies. We obtained four performances from the two participants and used these four times each without performing any averaging so as to have 16 instances in total. For the key velocity profiles, we followed the same procedure as we determined tempi and three performances out of the 31 valid baseline performances were randomly selected and the key velocity profiles of the three performances were averaged for each key (*Fig S3.1 I*). This is how we created 16 instances for three performance features (i.e., tempo, articulation and dynamics). We generated the 16 recordings where neither articulation nor dynamics was implemented (i.e., none) by combining the three performance features made from the baseline performances (*Fig S3.1 II*).

In order to have the other three types of recordings (i.e., both, art-only, dyn-only), we needed to create 16 instances for the sound duration when articulation was implemented and 16 instances for the key velocity profiles when dynamics was implemented. We followed the same procedure of averaging as we did to generate none recordings but we used articulation performances to determine the sound duration and dynamics performances to determine the key velocity profiles. To create art-only recordings, we replaced the articulation performance feature of none recordings with the one generated from articulation performances. Similarly for dyn-only recordings, we replaced the dynamics performance feature of none recordings with the one generated from dynamics performances. Lastly, to create both recordings, we replaced the articulation and dynamics performance features of none recordings with the ones generated from articulation and dynamics performances (*Fig S3.1 II*). Thus, we generated 16 recordings for each type.

Selection of four instances per type. After we generated 16 recordings for each type, we asked 3 musician raters including the first author to listen to each generated recording while sheet music (*Fig 3.1*) was displayed and rate to what extent each notated expression (i.e., articulation, dynamics) was implemented between 0 (not at all) and 5 (fully). Based on these ratings, we selected four recordings, which represented the characteristics of each type (i.e., both, art-only, dyn-only, none) the most (*Fig S3.1 III*). In order to fix the temporal feature of performance for four selected instances in each type, we randomly selected four tempo structures from 16 tempo instances and then replaced the tempo performance features of each recording type with the ones we selected (*Fig S3.2*).

Tweaking selected recordings. For dyn-only and none recordings, where notated articulation was not implemented, the recordings sounded too staccato. In terms of KOT, we calculated the difference between the mean of each recording for dyn-only and none recordings and the grand mean of art-only and both recordings and added that difference to all KOT values for dyn_only and none recordings, so that the KOT values of dyn-only and none recordings should shift closer to the grand mean of art_only and both recordings (*Fig S3.3*).

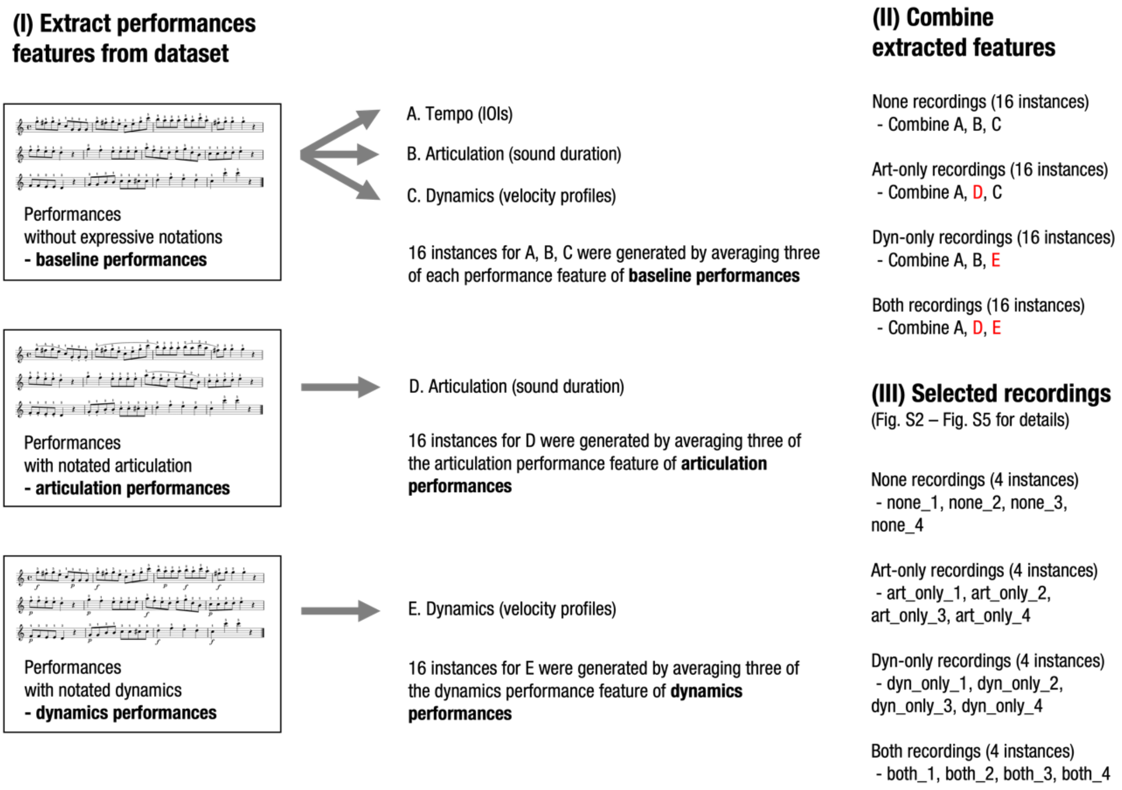

Figure S3.1. Schematic representation of the stimuli generation process.

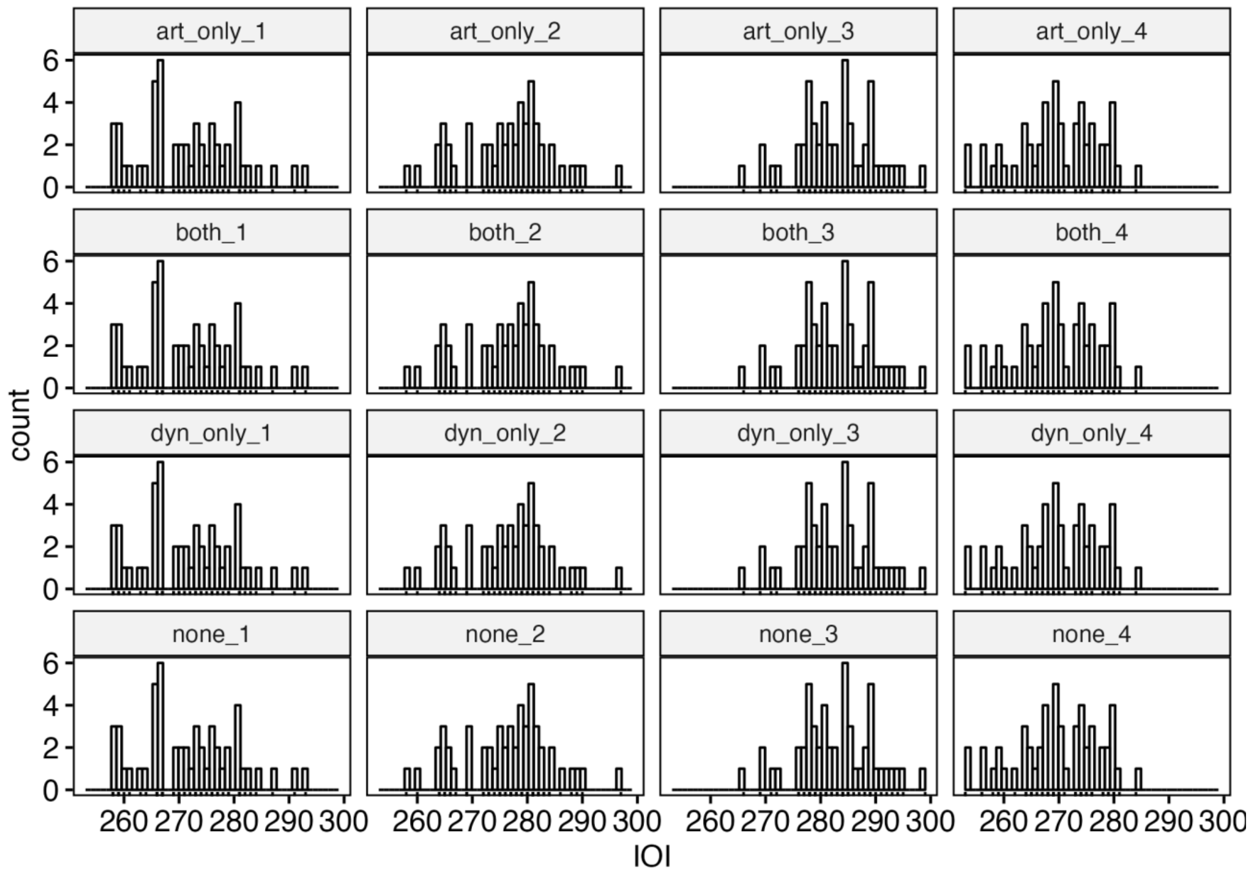

Figure S3.2. Histogram of interonset intervals (IOIs) for each recording. Since we fixed the temporal features across four instances, IOI profiles were identical for each type.

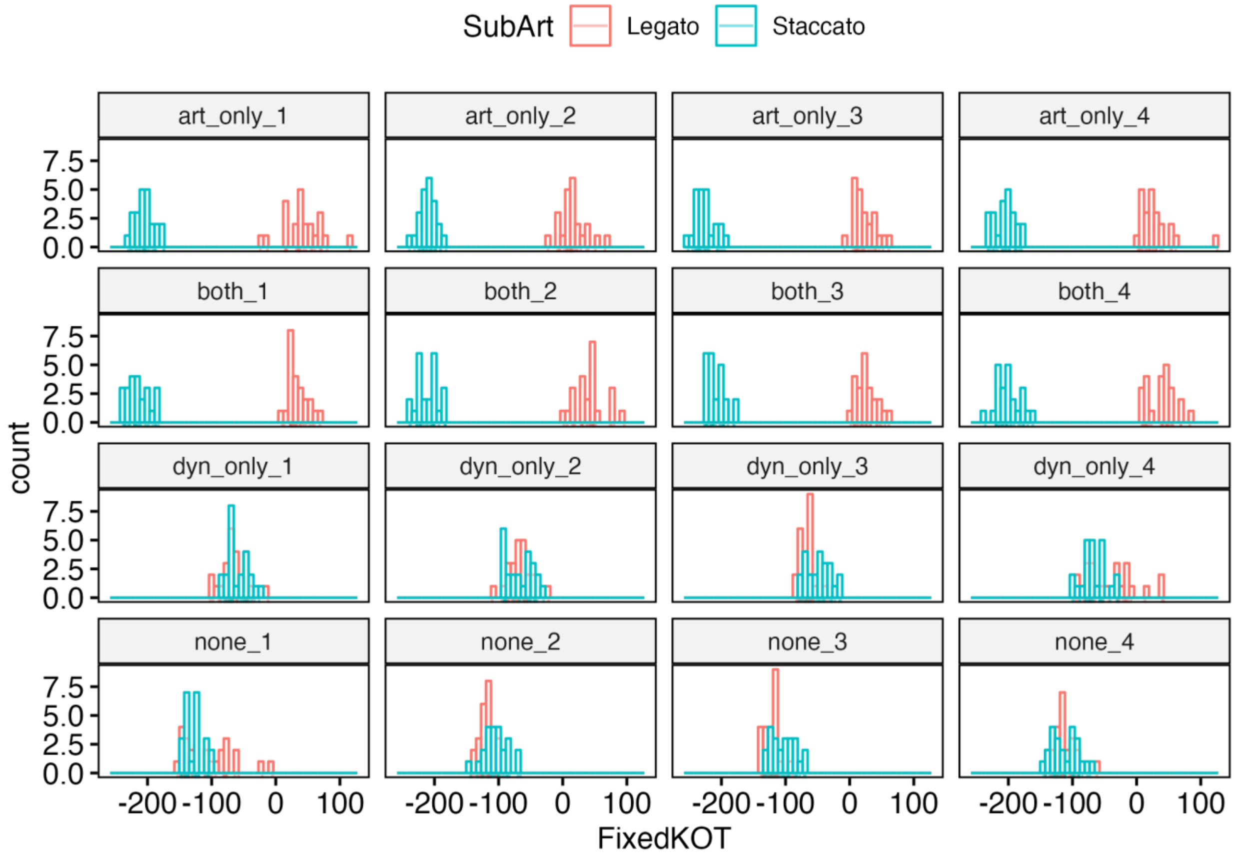

Figure S3.3. Histogram of key-overlap time (KOT) for each recording. For art-only and both recordings, there is a bimodal distribution, which represents the implementation of legato and staccato. For dyn-only and none recordings, there is no bimodal distribution (no difference between legato and staccato parts of performance). Red bars indicate KOT values where legato is notated. Blue bars indicate KOT values where staccato is notated.

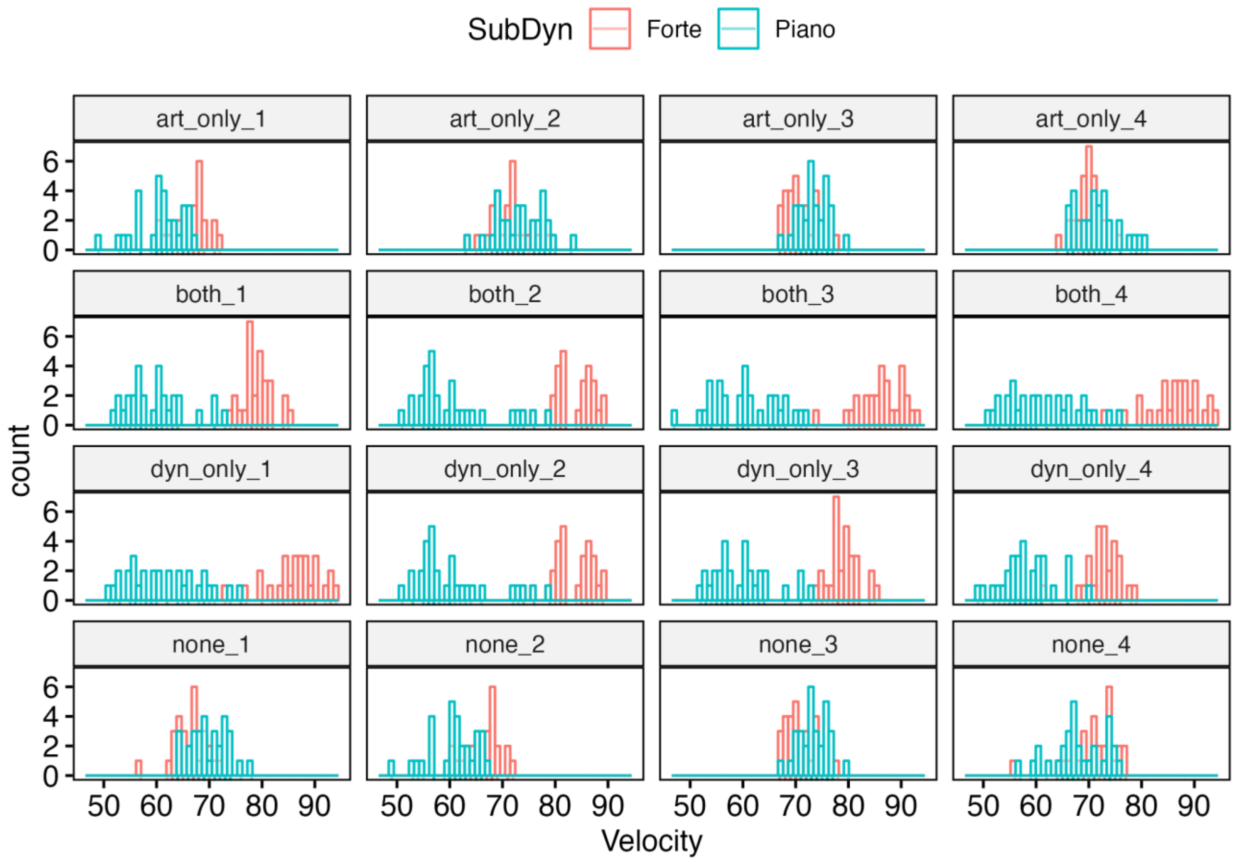

Figure S3.4. Histogram of key velocity (KV) for each recording. For dyn-only and both recordings, there is a bimodal distribution, which represents the implementation of forte and piano. For art-only and none recordings, there is no bimodal distribution (no difference between forte and piano parts of performance). Red bars indicate KV values where forte is notated. Blue bars indicate KV values where piano is notated.

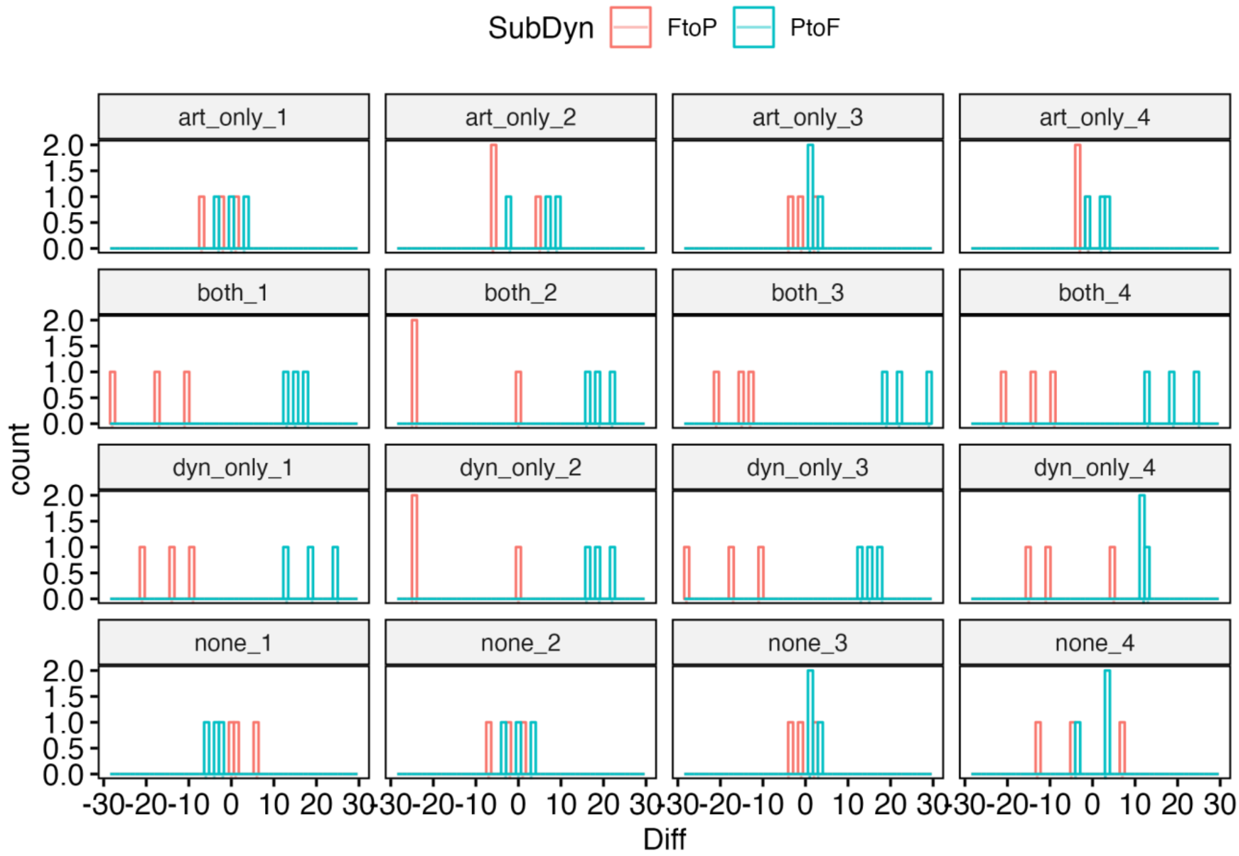

Figure S3.5. Histogram of KV difference for each recording. For dyn-only and both recordings, there is a bimodal distribution, which represents the transition from forte to piano or from piano to forte. For art-only and none recordings, there is no bimodal distribution (no contrast difference in transition parts between forte and piano). Red bars indicate KV difference of the transition from forte to piano. Blue bars indicate KV difference of the transition from piano to forte.

2: Comparison with baseline performance

To recall the instruction for the baseline performance, we asked participants to perform the piece expressively with their interpretation and to do their best as a performer. By comparing with the baseline performance, we can investigate how pedagogical intentions influence participants' performances.

KOT

We compared the baseline performance with the performances in response to recordings where dynamics was implemented to examine whether participants performed differently depending on whether articulation was implemented or not in each recording. We categorised performances into three groups (baseline, both (i.e., articulation-present, dynamics-present), dynamics-only (i.e., articulation-absent, dynamics-present)) and treated them as a factor Category.

Legato. There was a significant main effect of Category (F(1.45, 27.53) = 11.82, p < 0.001, η_G^2 = 0.060; Greenhouse-Geisser corrected). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment showed that there were differences between baseline and the other two categories (baseline and both: p < .001, baseline and dynamics-only: p = .008), suggesting that participants produced longer legato when they had the intention to teach (*Fig S3.6*, left).

Staccato. There was no significant main effect of Category (F(1.12, 21.28) = 0.75, p = 0.41, η_G^2 = 0.011; Greenhouse-Geisser corrected), suggesting that participants did not play staccato differently depending on whether they had the intention to teach (*Fig S3.6*, right).

KV

We compared the baseline performance with the performances in response to recordings where articulation was implemented to examine whether participants performed differently depending on whether dynamics was implemented or not in each recording. We categorised performances into three groups (baseline, both (i.e., articulation-present, dynamics-present), articulation-only (i.e., articulation-present, dynamics-absent)) and treated them as a factor Category.

Forte. There was no significant main effect of Category ($F(1.59, 30.18) = 1.43$, $p = 0.25$, $\eta_G^2 = 0.004$; Greenhouse-Geisser corrected), suggesting that participants did not play forte differently depending on whether they had the intention to teach (*Fig S3.7*, left).

Piano. There was a significant main effect of Category ($F(1.79, 34.09) = 5.68$, $p = 0.009$, $\eta_G^2 = 0.072$; Greenhouse-Geisser corrected). Post-hoc comparisons based on the estimated marginal means with Tukey adjustment showed that there were differences between baseline and the other two categories (baseline and both: $p = .05$, baseline and articulation-only: $p = .03$), suggesting that participants produced softer piano when they had the intention to teach (*Fig S3.7*, right).

KV Difference

We compared the baseline performance with the performances in response to recordings where articulation was implemented to examine whether participants performed differently depending on whether dynamics was implemented or not in each recording. We categorised performances into three groups (baseline, both (i.e., articulation-present, dynamics-present), articulation-only (i.e., articulation-present, dynamics-absent)) and treated them as a factor Category.

Forte to Piano. There was no significant main effect of Category (F(1.53, 29.15) = 1.73, p = 0.20, η_G^2 = 0.014; Greenhouse-Geisser corrected).

Piano to Forte. There was no significant main effect of Category (F(1.28, 24.39) = 2.27, p = 0.14, η_G^2 = 0.013; Greenhouse-Geisser corrected).

These results indicated that participants did not make dynamics contrast between forte and piano differently depending on whether they had the intention to teach (*Fig S3.8*).

IOIs

We compared the baseline performance with the performances in response to recordings where either both articulation and dynamics were implemented or neither of them was implemented. The reason for choosing this comparison was that we wanted to test whether having the intention to teach would lead to slower performances, either regardless of the learner's skill or especially when skills appeared to be lacking. We categorised performances into three groups (baseline, both (i.e., articulation-present, dynamics-present), none (i.e., articulation-absent, dynamics-absent)) and treated them as a factor Category.

There was no significant main effect of Category (F(1.11, 21.18) = 0.58, p = 0.47, η_G^2 = 0.012; Greenhouse-Geisser corrected), suggesting that participants kept the same tempo regardless of whether they had the intention to teach (*Fig S3.9*).

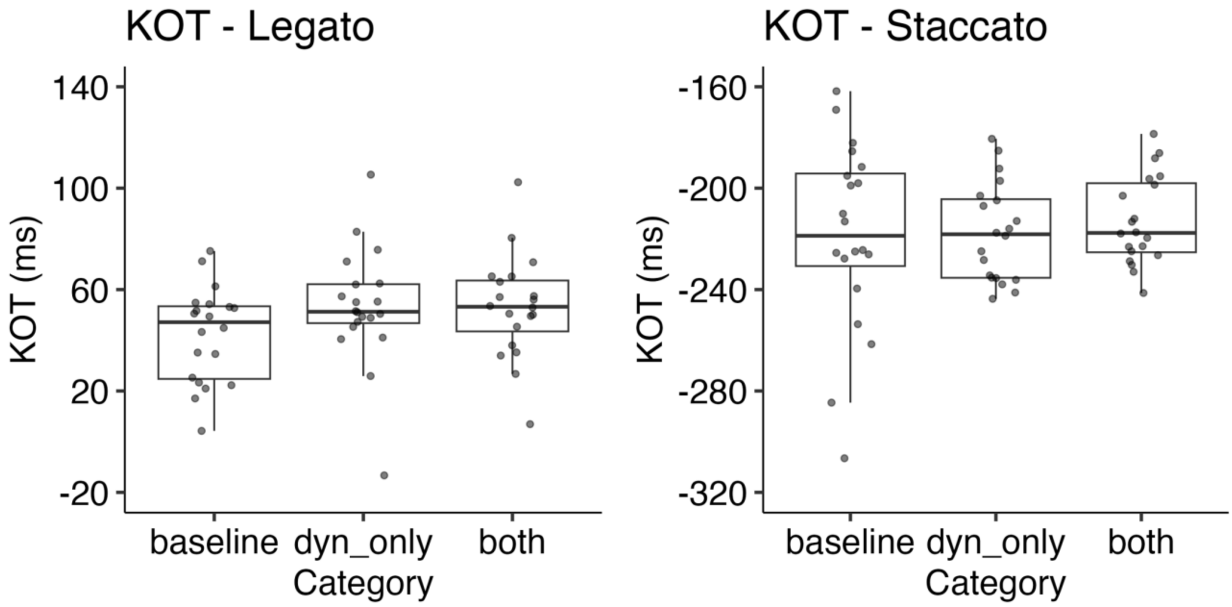

Figure S3.6. Comparison with the baseline performance in terms of KOT(ms) for each subcomponent; legato (left) and staccato (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

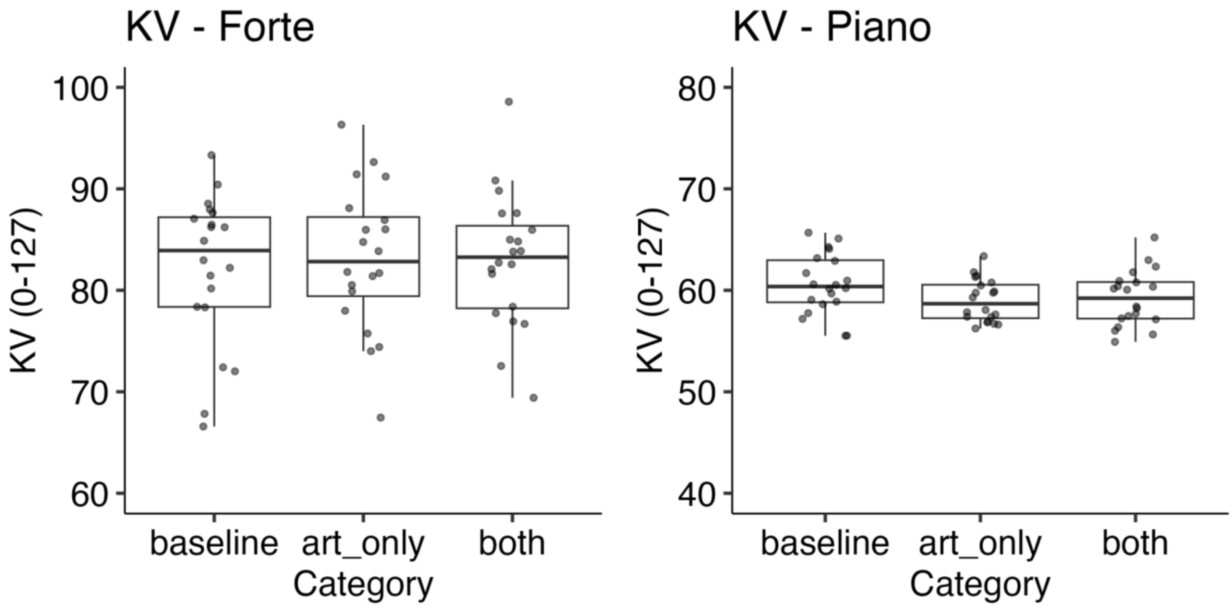

Figure S3.7. Comparison with the baseline performance in terms of KV (0-127) for each subcomponent; forte (left) and piano (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

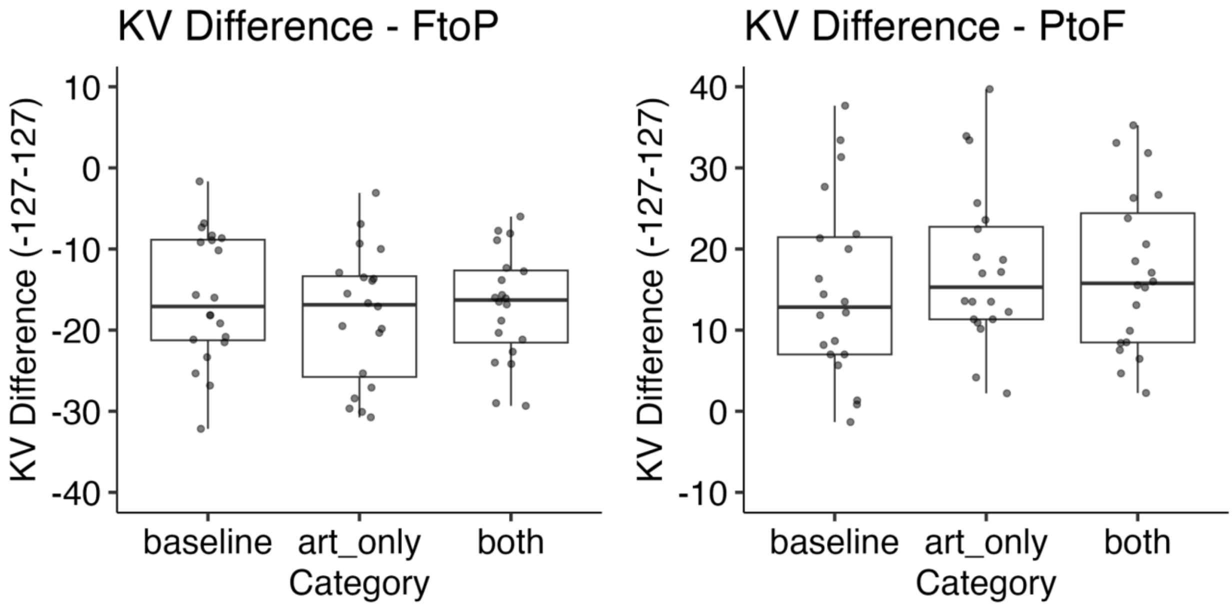

Figure S3.8. Comparison with the baseline performance in terms of KV Difference (-127-127) for each subcomponent; forte to piano (left) and piano to forte (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

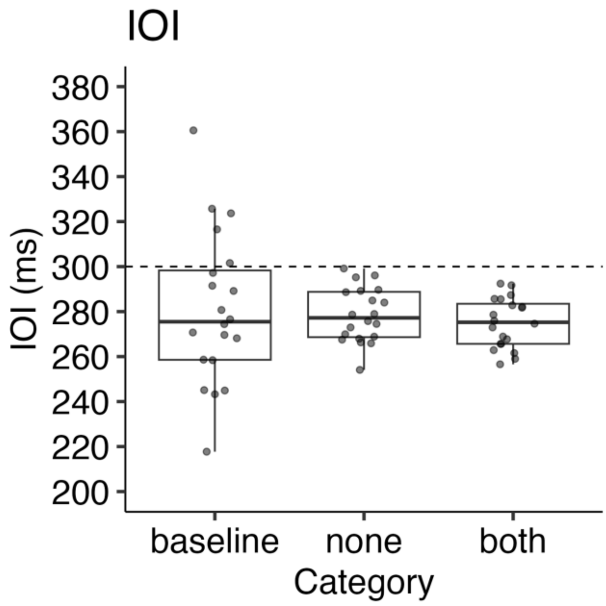

Figure S3.9. Comparison with the baseline performance in terms of IOIs (ms). A dashed line represents the tempo given by a metronome. Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

Chapter 4. Perception of exaggerated sound

4.1 Introduction

Learning from others is one of the important elements of skill acquisition. Not only are we able to learn by observing and imitating others, but also we benefit greatly from interacting with others such as teachers and peers (Tomasello et al., 1993). Adults are often being pedagogical to children to explain and transmit cultural conventions (Csibra & Gergely, 2009). Active teaching seems to play a crucial role not only to transit skills over generations but also to further develop sophisticated cultures, which cannot be achieved by one single individual or generation (Tennie et al., 2009). From a learner's perspective, it is important to identify informative teachers and infer teachers' expectations so that learners can acquire skills through interacting with teachers (Gweon, 2020; Veissière, Constant, Ramstead, Friston, & Kirmayer, 2020).

In pedagogical settings where teachers are supposed to convey useful information to learners, it has been found that teachers often modulate their behaviour for teaching purposes. For example, adults are likely to modulate their speech and action for infants to help them acquire skills (e.g., Brand et al., 2002; Saint-Georges et al., 2013). Further studies have revealed that even towards adult learners, people modulate their speech and action in a similar way as they did for infants (McEllin et al., 2017; Uther et al., 2007). Moreover, McEllin et al. (2018) demonstrated that people could identify informative intentions such as acting with others or teaching by relying on specific kinematics cues (e.g., velocity profiles of movements). These findings suggest that experts modified their speech and action to send teaching intentions and that novices could successfully perceive the intentions.

Tominaga et al. (2022) extended this line of research to expertise transmission where skills to be acquired are complex, such as when learning to perform with artistic expression. We investigated whether and how expert pianists communicatively modulate their performance when intending to teach the musical expressive techniques of articulation (the smoothness of sound) and dynamics (the loudness of sound). The results demonstrated that expert pianists systematically modulated their performance by playing slower or by exaggerating relevant aspects of the performance (e.g., producing shorter staccato or producing a larger contrast between forte and piano) when intending to teach musical expressive techniques. Therefore, it seems that experts exhibit pedagogical behaviours to highlight crucial performance aspects for potential learners.

Here we investigated whether the modulations that expert pianists make when they intend to teach are perceived by listeners as conveying teaching intentions. We started from the assumption that listeners would be able to infer intentions from listening to recordings because previous research revealed that sound alone is sufficient to communicate different aspects of a musical performance. For instance, listeners are generally able to infer performers' intended emotions when listening to recorded performances and musically skilled listeners are even better at decoding emotions from performances (e.g., Akkermans et al., 2019; Gabrielsson & Juslin, 1996).

In the present study, musicians listened to piano recordings where a musical expressive technique of either articulation or dynamics was implemented. They were asked to judge whether each recording was produced for teaching purposes or not. Half of the recordings was produced when pianists were instructed to play as if they were teaching the designated musical technique in a lesson (i.e., teaching recordings). The other half was produced when pianists were instructed to play as if they were performing it in a concert (i.e., performing recordings). First, we calculated the accuracy of participants' judgments to examine whether they could distinguish teaching recordings from performing recordings. Furthermore, recordings were quantified with regard to tempo, articulation and dynamics. Using correlations and multiple regression analysis, we examined which modulations of performance were used to infer teaching intentions.

If performers' intentions are understood by learners with basic musical skills, they should be able to distinguish teaching recordings from performing recordings. Moreover, the same modulations of pedagogical performance parameters observed in our previous experiments such as slower demonstration and exaggerated performance, should be used to infer teaching intentions by the learners. In Experiment 1, participants listened to performances that followed simple musical scales. In Experiment 2, participants listened to performances of a more complex piece.

4.2 Experiment 1

4.2.1 Methods

Participants We recruited 21 participants who had at least six years of training in any musical instrument. They were able to read sheet music and knew two musical expressive techniques of articulation and dynamics. One participant was excluded due to an experimental error. Therefore, 20 participants (13 female) were included for data analysis and had 11.8 years of musical training on average (*SD* = 5.62). They were all right-handed with a mean age of 28.8 (*SD* = 9.09). All participants were recruited through an online participant platform (SONA system, https://www.sona-systems.com). The study (No. 2020/02) was approved by the Psychological Research Ethics Board (PREBO) CEU PU in Austria.

Apparatus The experiment was programmed in Python 3.8.2 using the PsychoPy Python library (2020.2.4; https://www.psychopy.org/) on a MacBook Pro with Mac OS X Catalina 10.15.6. Stimuli were played using the Mido Python library (1.2.9; https://mido.readthedocs.io/en/latest/) on a Max/MSP patcher (8.1.7; https://cycling74.com/products/max). During the experiment, participants listened to the stimuli via headphones (Audio-Technica ATH-M50X).

Stimuli We selected stimuli from the performances obtained in our previous experiments (Tominaga et al., 2022). Stimuli were produced by actual pianists on a weighed Yamaha MIDI (Musical Instrument Digital Interface) digital piano and recorded as MIDI files. Multiple pianists played one piece of music with a musical expressive technique of either articulation (*Fig 4.1 A*) or dynamics (*Fig 4.1 B*). Articulation refers to the smoothness of sound, which is comprised of legato and staccato. Legato indicates smooth and connected sound whereas staccato indicates sharp and separate sound. Dynamics refers to the loudness of sound, which is comprised of forte and piano. Forte indicates loud sound while piano indicates soft sound. The piece was taken from "A Dozen a Day - Play with Ease in Many Keys" by Edna-Mae Burnam and modified for the experiment. The stimuli were performed at around 80 quarter-beats per minute.

In Tominaga et al. (2022), participants were asked to perform the piece with either articulation or dynamics in two different conditions. In the teaching condition, participants were instructed to perform the piece with the designated expressive technique as if they were teaching it to students (e.g., in a lesson). In the performing condition, participants were instructed to perform the piece with the designated expressive technique as if they were performing it to an audience (e.g., in a concert). In Tominaga et al. (2022; Experiment 1), there were 453 valid performances (i.e., performances without any pitch errors) from the teaching condition and 436 valid performances from the performing condition. For the current experiment, 96 recordings were chosen from the valid performances. We randomly sampled 24 articulation recordings and 24 dynamics recordings from the teaching condition as well as 24 articulation recordings and 24 dynamics recordings from the performing condition. It is important to note that the recordings from the teaching condition did not necessarily exhibit specific features of teaching performance that we found in our previous experiments (e.g., exaggeration) since we randomly sampled the performances from multiple pianists.

Procedure Upon arrival, participants read the information sheet about the experiment and gave informed consent prior to participation. In the experiment, all instructions were displayed on a computer screen in front of the participants and an experimenter also explained the procedure. Participants were instructed that they were going to listen to piano recordings with one musical expressive technique of either articulation or dynamics, which were either produced as if a pianist were teaching the designated expressive technique to students (e.g., in a lesson) or as if a pianist were performing it to an audience (e.g., in a concert). In each trial, participants listened to one recording and were asked whether the recording was produced for teaching purposes or not. Participants responded by pressing either a yes (left arrow key) or no (right arrow key) button. While listening to each recording, sheet music, which corresponded to the recording, was shown on the screen in front of the participants (*Fig 4.2*). Participants were allowed to listen to each recording only once.

There were two blocks and each block only included the recordings with one musical expressive technique of either articulation or dynamics. Each block consisted of four practice trials and 48 experimental trials. Each recording was evaluated only once in the experiment.

All participants completed both blocks. The order of the blocks was counterbalanced across participants. The order of the recordings was randomised within each block.

At the end of the experiment, participants filled in a questionnaire about their demographic information and their prior experience with playing musical instruments.

4.2.2 Data analysis

Data processing and statistical analysis were performed in R version 4.0.5. Correlation analysis was performed with the standard *cor* function and regression models for multiple regression were fit with the standard *lm* function from the *stats* R package. Stimuli (MIDI files) were converted to numerical data in terms of time, pitch and velocity for the onset and offset of each note using the *tuneR* R package (https://cran.r-project.org/web/packages/tuneR/tuneR.pdf).

Accuracy. First, we examined whether participants could accurately recognise the stimuli chosen from the teaching condition in Tominaga et al., (2022; Experiment 1) as teaching and the stimuli chosen from the performing condition in Tominaga et al., (2022; Experiment 1) as performing. We compared how accurate participants were against the chance level (50%). Correct responses were either pressing the yes button when listening to teaching recordings or pressing the no button when listening to performing recordings. Incorrect responses were either pressing the yes button when listening to performing recordings or pressing the no button when listening to teaching recordings.

Correlations and multiple regression. Stimuli were quantified regarding tempo (interonset intervals; IOIs), articulation (key-overlap time; KOT), dynamics (key velocity; KV) and dynamics contrast (key velocity difference; KV-Diff) only for 16th notes. Interonset intervals are time intervals between onsets of adjacent notes. Larger IOIs indicate slower tempo while smaller IOIs indicate faster tempo. Key-overlap time is the time overlap between two adjacent notes, namely the difference between the offset time of the current note and the onset time of the ensuing note (e.g., Bresin & Battel, 2000). Positive KOT values indicate a legato style whereas negative KOT values indicate a staccato style. Key velocity is obtained from MIDI data to describe how fast a performer hit the key. Larger KV values indicate a forte style while smaller KV values indicate a piano style. Additionally, we also measured dynamics contrast where one subcomponent of the technique moves to the other (e.g., from forte to piano, from staccato to legato) to illustrate how much dynamics contrast a performer made at transition points.

To investigate the relationships between performance features (i.e., IOIs, KOT, KV, KV-Diff) and participants' judgments as teaching (i.e., what percentage of participants responded as "yes"), we first performed separate correlation analyses for each performance feature. Second, we run multiple regression that included all four performance features to examine the strongest predictor of participants' judgments as teaching. Since articulation and dynamics were comprised of two opposite directional values (i.e., legato vs. staccato, forte vs. piano), we created four separate models that included different parts of the piece. Each model considered only parts of the piece that comprised either legato, staccato, forte or piano. For articulation recordings, there were two models. The Legato model (*Fig 4.1* (1)) considered only legato parts of the performances. We entered the legato parts of KOT and KV and KV-Diff or transition points from legato to staccato. The Staccato model (*Fig 4.1* (2)) considered only staccato parts of the performances. We entered the staccato parts of KOT and KV and KV-Diff of transition points from staccato to legato. Similarly, there were two models for dynamics recordings. The Forte model (*Fig 4.1* (3)) considered only forte parts of the performances for KV and KOT, and transition points from forte to piano for KV-Diff. The Piano model (*Fig 4.1* (4)) considered only piano parts of the performances for KV and KOT, and transition points from piano to forte for KV-Diff. With regard to tempo (IOIs), there was only one value for each recording regardless of the subcomponents because tempo was consistent across the performance. Therefore, we entered the same tempo value for the Legato and Staccato models or the Forte and Piano models.

Figure 4.1. Stimuli. (A)Articulation. The curved line (slur) indicates legato and the dots indicate staccato. In multiple regression analysis, (1) corresponds to Legato model and (2) corresponds to Staccato model. (B)Dynamics. The symbol 'f' denotes forte and the symbol 'p' denotes piano. In multiple regression analysis, (3) corresponds to Forte model and (4) corresponds to Piano model. Only the parts composed of 16th notes (i.e., (1), (2), (3), (4)) were used for data analysis.

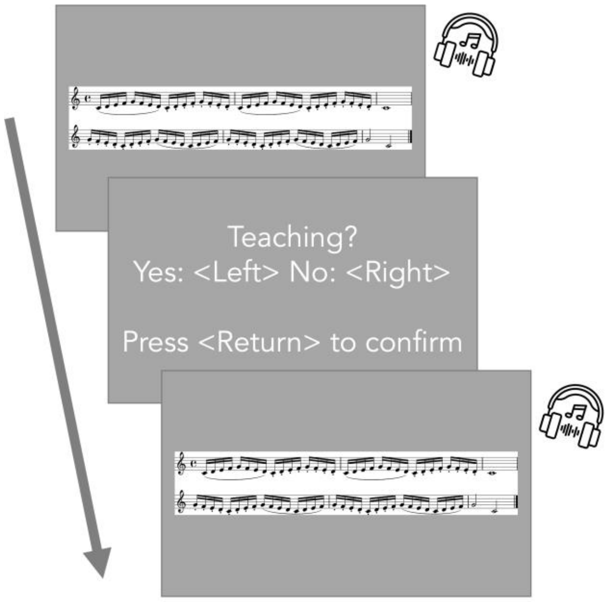

Figure 4.2. Procedure. Participants listened to a recording via headphones while corresponding sheet music was displayed on a monitor. They were required to respond by pressing the left-arrow (yes) or right-arrow (no) key for each judgment. Headphone image from Flaticon.com.

4.2.3 Results

All results were reported as significant at $p < 0.05$.

Accuracy

A one sample *t*-test was performed to compare the accuracy of participants' judgments against chance level (50%). The mean percentage of correct answers [$M = 52.7$, $SD = 4.90$] was significantly higher than chance ($t(19) = 2.47$, $p = .02$, cohen's $d = 0.55$).

We also performed a one sample *t*-test for each technique separately. For articulation, the mean percentage of correct answers [$M = 50.2$, $SD = 6.55$] was not significantly different from chance ($t(19) = 0.14$, $p = .89$, cohen's $d = 0.03$). For dynamics, the mean percentage of correct answers [$M = 55.2$, $SD = 7.81$] was significantly higher than chance ($t(19) = 2.98$, $p =$

.01, cohen's *d* = 0.67). A paired *t*-test revealed that there was a significant difference between the two techniques in terms of the accuracy (*t*(19) = -2.12, *p* = .05, cohen's *d* = -0.69), suggesting participants chose correct answers more for dynamics recordings than articulation recordings.

Correlations

Tempo (IOIs). Performance tempi (IOIs) were significantly correlated with participants' judgments as teaching for both techniques (Articulation; *r(46)* = .77, *p* < .001, Dynamics; *r(46)* = .42, *p* = .003, *Fig 4.3*). Participants identified slower performances as teaching.

Articulation (KOT). For articulation recordings, there was a significant relationship between KOT values and participants' judgments as teaching (*Fig 4.4*, left). Specifically, performances with shorter staccato (*r(46)* = -.73, *p* < .001) and longer legato (*r(46)* = .40, *p* = .005) were more likely to be judged as teaching.

For dynamics recordings, there was no significant relationship between KOT values and participants' judgments as teaching (Forte; *r(46)* = -.04, *p* = .81, Piano; *r(46)* = -.19, *p* = .19, *Fig 4.4*, right).

Dynamics (KV). For dynamics recordings, there was a significant relationship between KV values and participants' judgments as teaching (*Fig 4.5*, right). Specifically, performances with louder forte were more likely to be judged as teaching (*r(46)* = .45, *p* = .001). However, there was no significant relationship between KV values for piano and participants' judgments as teaching (*r(46)* = -.22, *p* = .13).

For articulation recordings, there was no significant relationship between KV values and participants' judgments as teaching (Legato; $r(46) = .10$, $p = .52$, Staccato; $r(46) = -.02$, $p = .87$, *Fig 4.5*, left).

Dynamics contrast (KV-Diff). For dynamics recordings, there was a significant relationship between KV difference between forte and piano and participants' judgments as teaching (*Fig 4.6*, right). Specifically, performances with larger contrasts between forte and piano were more likely to be judged as teaching (From Forte to Piano; $r(46) = -.50$, $p < .001$, From Piano to Forte; $r(46) = .62$, $p < .001$).

For articulation recordings, there was no significant relationship between KV difference between legato and staccato and participants' judgments as teaching (From Legato to Staccato; $r(46) = -.26$, $p = .07$, From Staccato to Legato; $r(46) = -.19$, $p = .21$, *Fig 4.6*, left).

Multiple regression

In order to further investigate which feature of performance contributed the most to participants' judgments as teaching, multiple regression analyses were conducted. Statistical model assumptions were tested using the *performance* R package (Lüdecke, Ben-Shachar, Patil, Waggoner, & Makowski, 2021). Since articulation and dynamics consisted of two opposite subcomponents (i.e., legato vs. staccato, forte vs. piano) and therefore cannot be summed up to represent each technique as one value, we reported four separate regression models for each subcomponent (see details in *Data analysis* and *Fig 4.1*).

Legato. A multiple regression analysis was conducted to predict participants' judgments as teaching based on performance features of tempo (IOIs), articulation (KOT for legato parts), dynamics (KV for legato parts) and dynamics contrast (KV-Diff from legato to staccato). The result of the regression indicated that the model explained 64.6 % of the variance ($F(4, 43) = 22.5$, $p < .001$). It was found that tempo (IOIs; $\beta = 0.78$, $p < .001$) and articulation for the legato parts (KOT; $\beta = 0.26$, $p = .004$) were significant predictors of participants' judgments as teaching.

Staccato. A multiple regression analysis was conducted to predict participants' judgments as teaching based on performance features of tempo (IOIs), articulation (KOT for staccato parts), dynamics (KV for staccato parts) and dynamics contrast (KV-Diff from staccato to legato). The result of the regression indicated that the model explained 64.0 % of the variance ($F(4, 43) = 21.9$, $p < .001$). It was found that tempo (IOIs; $\beta = 0.52$, $p = .002$) and articulation for the staccato parts (KOT; $\beta = -0.28$, $p = .02$) were significant predictors of participants' judgments as teaching.

Forte. A multiple regression analysis was conducted to predict participants' judgments as teaching based on performance features of tempo (IOIs), articulation (KOT for forte parts), dynamics (KV for forte parts) and dynamics contrast (KV-Diff from forte to piano). The result of the regression indicated that the model explained 35.9 % of the variance ($F(4, 43) = 7.58$, $p < .001$). It was found that tempo (IOIs; $\beta = 0.35$, $p = .007$) and dynamics for the forte parts (KV; $\beta = 0.73$, $p = .05$) were significant predictors of participants' judgments as teaching.

Piano. A multiple regression analysis was conducted to predict participants' judgments as teaching based on performance features of tempo (IOIs), articulation (KOT for piano parts), dynamics (KV for piano parts) and dynamics contrast (KV-Diff from piano to forte). The result of the regression indicated that the model explained 51.2 % of the variance ($F(4, 43) = 13.3$, $p < .001$). It was found that tempo (IOIs; $\beta = 0.38$, $p < .001$) and dynamics contrast from piano to forte (KV-Diff; β = 0.99, $p < .001$) were significant predictors of participants' judgments as teaching.

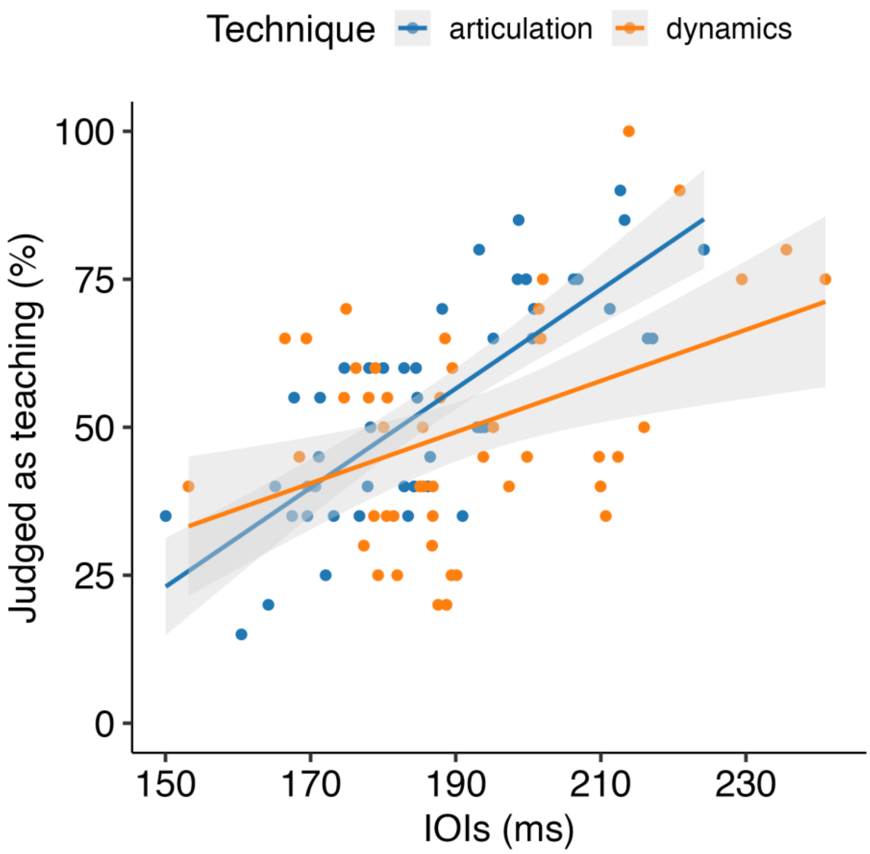

Figure 4.3. Experiment 1: Scatter plot showing the correlation between tempo feature (IOIs) and average participants' judgments as teaching for each recording. Therefore, each dot represents each stimulus. The grey bands display the 95 percent confidence intervals.

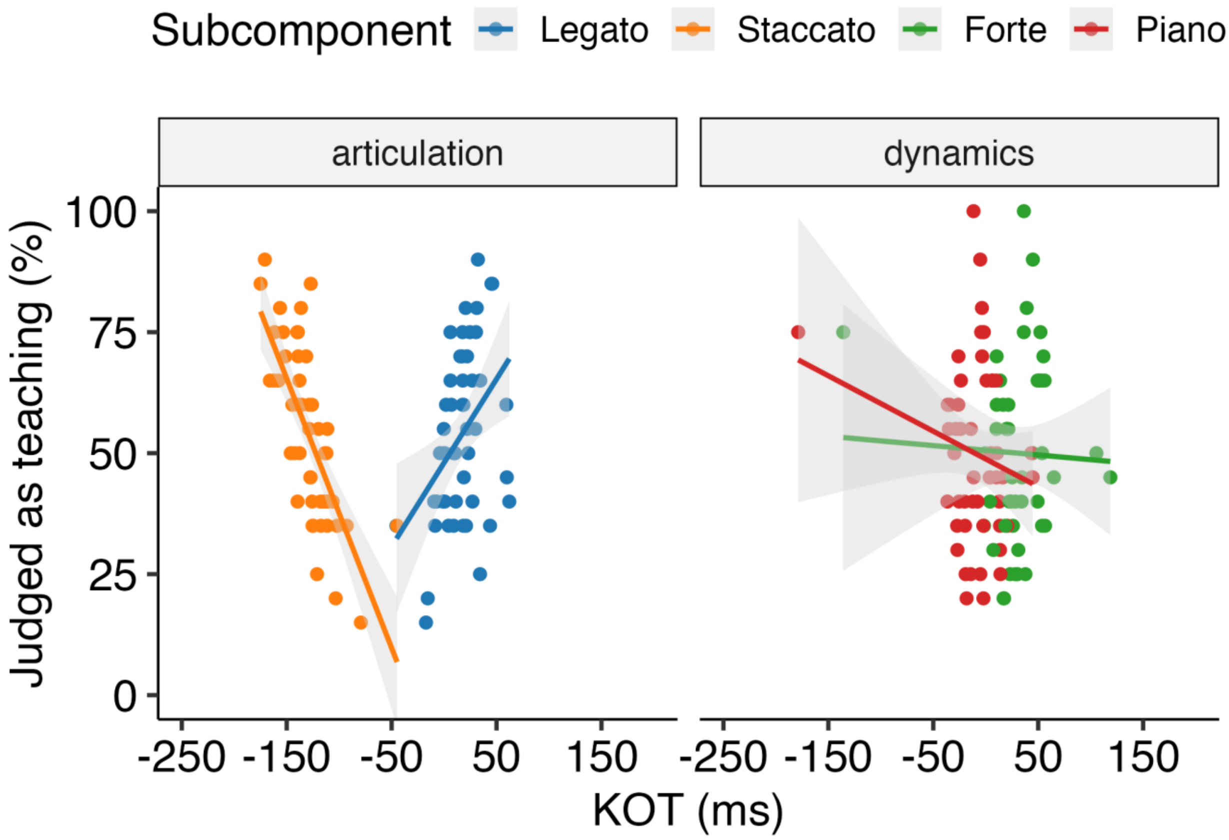

Figure 4.4. Experiment 1: Scatter plot showing the correlation between articulation feature (KOT) and average participants' judgments as teaching for each recording. Therefore, each dot represents each stimulus. The grey bands display the 95 percent confidence intervals.

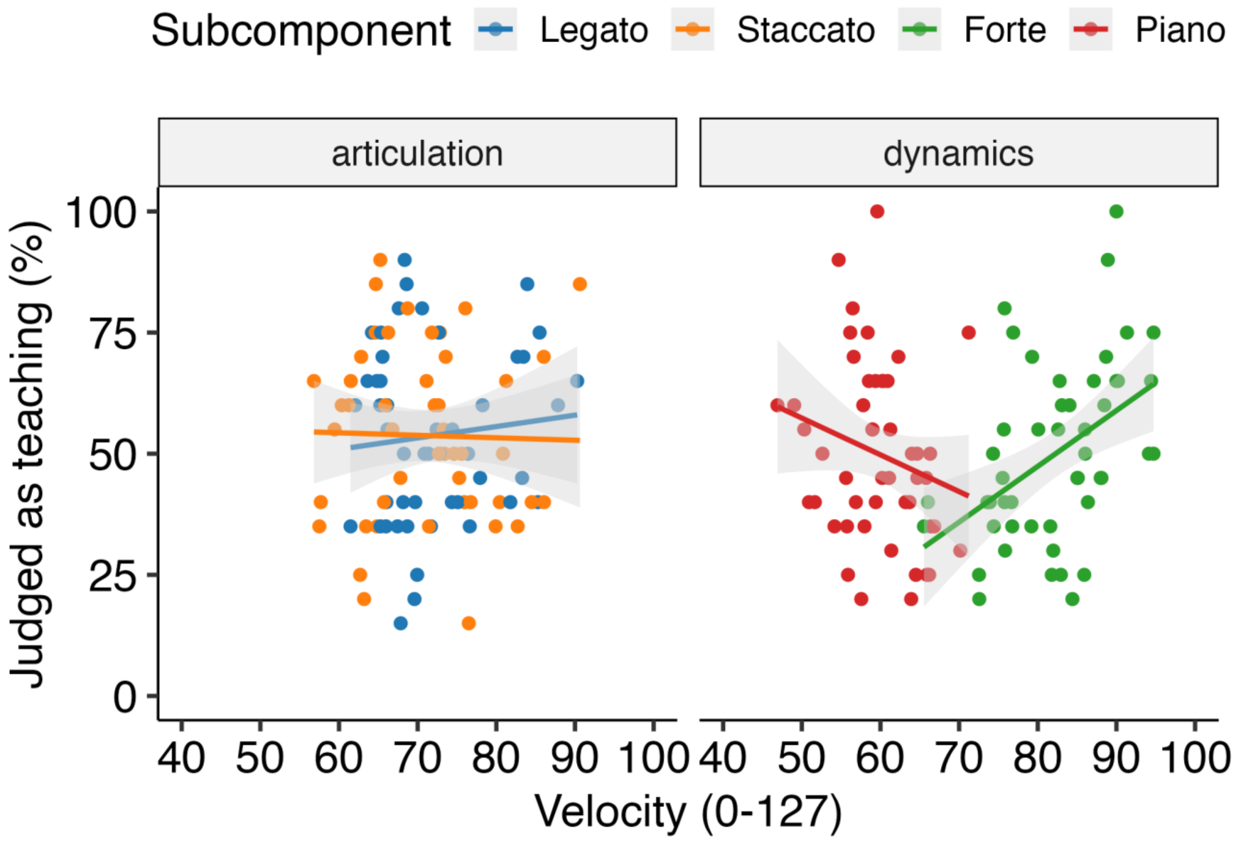

Figure 4.5. Experiment 1: Scatter plot showing the correlation between dynamics feature (KV) and average participants' judgments as teaching for each recording. Therefore, each dot represents each stimulus. The grey bands display the 95 percent confidence intervals.

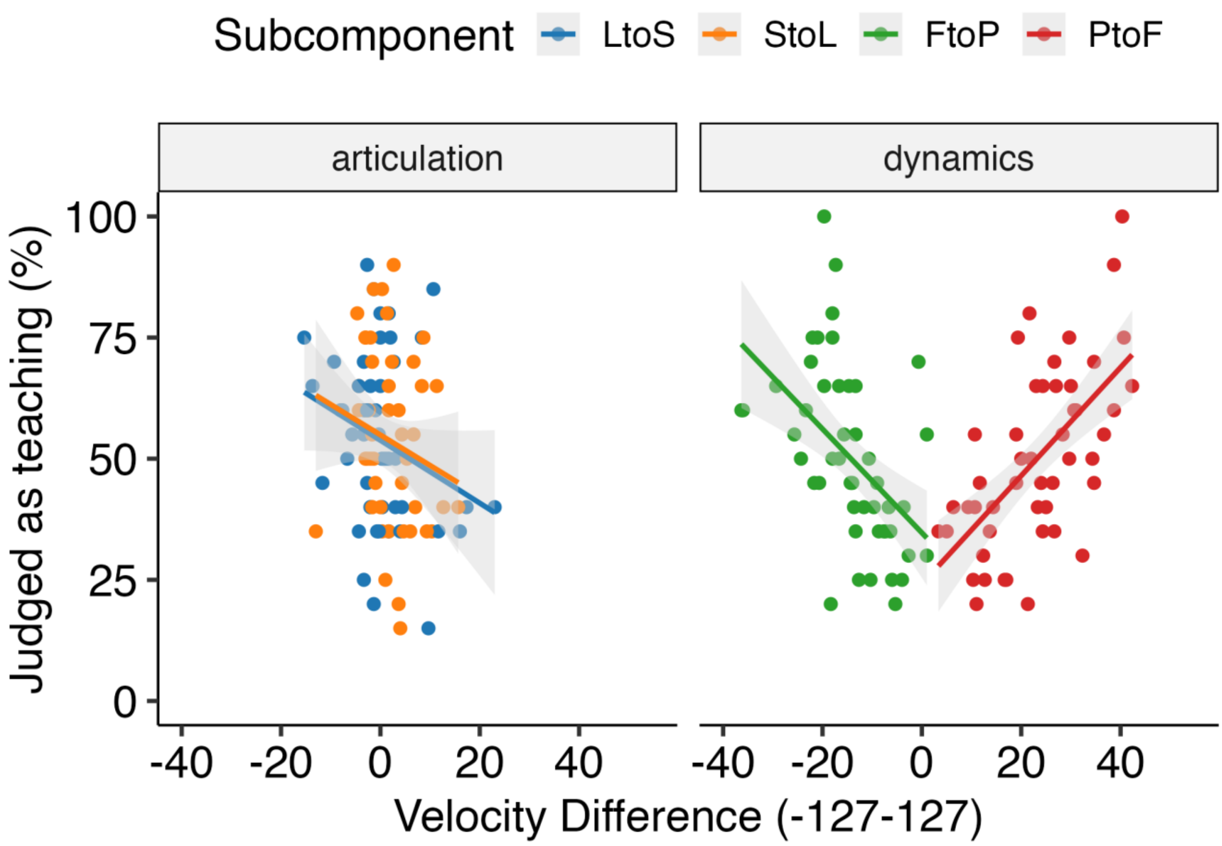

Figure 4.6. Experiment 1: Scatter plot showing the correlation between dynamics contrast feature (KV-Diff) and average participants' judgments as teaching for each recording. Therefore, each dot represents each stimulus. The grey bands display the 95 percent confidence intervals. LtoS; Legato to Staccato, StoL; Staccato to Legato, FtoP; Forte to Piano, PtoF; Piano to Forte.

4.2.4 Discussion

Experiment 1 investigated whether musicians could distinguish teaching recordings from performing recordings and which features of piano performance made them infer teaching intentions. The results demonstrated that musicians could choose correct answers above chance. Also, it was found that the accuracy of participants' judgments was better for dynamics recordings than for articulation recordings.

Performances with slower tempo were more likely to be judged as teaching by musicians regardless of which expressive technique was implemented in the piece. For articulation recordings, performances with longer legato and shorter staccato were tended to be judged as teaching. For dynamics recordings, louder performances were more likely to be judged as teaching whereas there was no relationship between softer sound and participants' judgments as teaching. Importantly, performances with larger contrasts between forte and piano for both directions (i.e., from forte to piano, from piano to forte) were more likely to be judged as teaching. This result may suggest that dynamics contrast might be reliably used to infer teaching intentions, rather than absolute dynamics values themselves. Moreover, multiple regression analyses implied that tempo feature was the strongest predictor of participants' judgments as teaching in general whereas there were specific predictors depending on which expressive technique was implemented in the piece. These performance features were overall consistent with what expert pianists did in our previous experiments for teaching purposes. Therefore, our findings suggest that musicians may rely on generic pedagogical behaviours (e.g., slower demonstration, exaggeration) to infer teaching intentions of expert pianists when listening to recorded performances.

4.3 Experiment 2

The aim of Experiment 2 was to replicate the findings in Experiment 1 with a more complex piece of music. Given the findings in Experiment 1, we predicted that slower performance would be likely to be judged as teaching regardless of which expressive technique (i.e., articulation or dynamics) was implemented. Also performances with exaggerated articulation and dynamics (in particular, longer legato and shorter staccato, larger contrasts between forte and piano) would be likely to be judged as teaching.

4.3.1 Methods

Participants We recruited 21 participants who had at least six years of training in any musical instrument or singing. They were able to read sheet music and knew two musical expressive techniques of articulation and dynamics. One participant was excluded because s/he did not understand the instructions. Therefore, 20 participants (10 female) were included for data analysis and had 12.65 years of training on average in any musical instrument or singing (*SD* = 5.40). Most people were right-handed (left; 1) with a mean age of 33.55 (*SD* = 12.80). As Experiment 1, all participants were recruited through the SONA system and the study (No. 2020/02) was approved by the PREBO CEU PU in Austria.

Apparatus and procedure The apparatus and procedure were identical to Experiment 1 except that each block consisted of four practice trials and 36 experimental trials. The number of trials was reduced due to the time constraint of the experiment.

Stimuli As Experiment 1, we selected stimuli from our previous experiments (Tominaga et al., 2022; Experiment 2). The excerpt was taken from "Sonatina Op.36 (No.3) in C major" by Muzio Clementi and modified for the experiment. The excerpt was performed with either articulation (*Fig 4.7 A*) or dynamics (*Fig 4.7 B*). The stimuli were performed around 100 - 120 quarter-beats per minute.

For the current experiment, 72 performances were chosen from the valid performances in Tominaga et al. (2022; Experiment 2). There were 248 valid performances in the teaching condition and 256 valid performances in the performing condition. We randomly sampled 18 articulation performances and 18 dynamics performances from the teaching condition as well as 18 articulation performances and 18 dynamics performances from the performing condition. Again, it is important to note that each performance from the teaching condition did not necessarily exhibit specific features of teaching that we found in the previous experiments (e.g., exaggeration) since we randomly sampled the performances.

Data analysis The data analysis was almost identical to Experiment 1. Only the 8th notes with expressive notations were included for data analysis. As a result, only one 8th note in the 4th measure without any expression was not included (see *Fig 4.7*).

Figure 4.7. Stimuli. (A)Articulation. The curved line (slur) indicates legato and the dots indicate staccato. In multiple regression analysis, (1) corresponds to Legato model and (2) corresponds to Staccato model. (B)Dynamics. The symbol 'f' denotes forte and the symbol 'p' denotes piano. In multiple regression analysis, (3) corresponds to Forte model and (4)

corresponds to Piano model. Only the 8th notes with expressive notations (i.e., (1), (2), (3), (4)) were used for data analysis.

4.3.2 Results

All results were reported as significant at $p < 0.05$.

Accuracy

A one sample *t*-test was performed to compare the accuracy of participants' judgments against chance level (50%). The mean percentage of correct answers [*M* = 52.8, *SD* = 4.78] was significantly higher than chance ($t(19) = 2.66$, $p = .02$, cohen's $d = 0.60$).

We also performed a one sample *t*-test for each technique separately. For articulation, the mean percentage of correct answers [*M* = 52.4, *SD* = 6.94] was not significantly different from chance ($t(19) = 1.52$, $p = .14$, cohen's $d = 0.34$). Also for dynamics, the mean percentage of correct answers [*M* = 53.3, *SD* = 8.53] was not significantly different from chance ($t(19) = 1.75$, $p = .10$, cohen's $d = 0.39$). A paired *t*-test revealed that there was no significant difference between the two techniques in terms of the accuracy ($t(19) = -0.35$, $p = .73$, cohen's $d = -0.13$).

Correlations

Tempo (IOIs). There was a significant relationship between performance tempi (IOIs) and participants' judgments as teaching only for dynamics recordings (Articulation; $r(34) = .25$, $p = .15$, Dynamics; $r(34) = .39$, $p = .02$, *Fig 4.8*). Participants tended to identify slower performances as teaching for dynamics recordings.

Articulation (KOT). For articulation recordings, there was no significant relationship between KOT values and participants' judgments as teaching (Legato; *r(34)* = -.03, *p* = .88, Staccato; *r(34)* = -.15, *p* = .39, *Fig 4.9*, left).

For dynamics recordings, there was no significant relationship between KOT values for forte and participants' judgments as teaching (*r(34)* = -.11, *p* = .52). However, there was a significant relationship between KOT values for piano and participants' judgments as teaching (*r(34)* = -.35, *p* = .03), suggesting that performances with staccato-style piano were more likely to be considered as teaching performance (*Fig 4.9*, right). However, this significance disappeared after removing one outlier outside of 2 standard deviations in terms of KOT values for piano.

Dynamics (KV). For dynamics recordings, there was a significant relationship between KV values and participants' judgments as teaching (*Fig 4.10*, right). Specifically, performances with louder forte (*r(34)* = .45, *p* = .007) and softer piano (*r(34)* = -.45, *p* = .006) were more likely to be judged as teaching.

For articulation recordings, there was no significant relationship between KV values and participants' judgments as teaching (Legato; *r(34)* = .08, *p* = .63, Staccato; *r(34)* = .21, *p* = .22, *Fig 4.10*, left).

Dynamics contrast (KV-Diff). For dynamics recordings, there was a significant relationship between KV difference between forte and piano and participants' judgments as teaching (*Fig 4.11*, right). Specifically, performances with larger contrasts between forte and piano were more likely to be judged as teaching (From Forte to Piano; *r(34)* = -.75, *p* < .001, From Piano to Forte; *r(34)* = .59, *p* < .001).

For articulation recordings, there was no significant relationship between KV difference between transition points from legato to staccato and participants' judgments as teaching (*r(34)* = .23, *p* = .18). However, there was a significant relationship between the transition points from staccato to legato and participants' judgments as teaching (*r(34)* = .36, *p* = .03), suggesting that performances with larger contrasts from staccato to legato were more likely to be considered as teaching performance (*Fig 4.11*, left).

Multiple regression

As Experiment 1, we performed multiple regression analyses to further investigate which feature of performance contributed the most to participants' judgments as teaching. Again, since articulation and dynamics consisted of two opposite subcomponents (i.e., legato vs. staccato, forte vs. piano) and therefore cannot be summed up to represent each technique as one value, we reported four separate regression models for each subcomponent (see details in *Data analysis* in Experiment 1 and *Fig 4.7*).

Legato. A multiple regression analysis was conducted to predict participants' judgments as teaching based on performance features of tempo (IOIs), articulation (KOT for legato parts), dynamics (KV for legato parts) and dynamics contrast (KV-Diff from legato to staccato). The result of the regression indicated that the model explained 19.8 % of the variance ($F(4, 31) = 3.16$, $p = .03$). It was found that tempo (IOIs; $\beta = 0.58$, $p = .01$), dynamics (KV; $\beta = 1.10$, $p = .02$) and dynamics contrast (KV-Diff; $\beta = 1.90$, $p = .006$) for the legato parts were articulation for the legato parts were significant predictors of participants' judgments as teaching. However, articulation (KOT; $\beta = -0.08$, $p = .47$) for the legato parts was not a significant predictor.

Staccato. A multiple regression analysis was conducted to predict participants' judgments as teaching based on performance features of tempo (IOIs), articulation (KOT for staccato parts), dynamics (KV for staccato parts) and dynamics contrast (KV-Diff from staccato to legato). The overall model was not statistically significant (R^2 = 12.3, $F(4, 31)$ = 2.22, p = .09).

Forte. A multiple regression analysis was conducted to predict participants' judgments as teaching based on performance features of tempo (IOIs), articulation (KOT for forte parts), dynamics (KV for forte parts) and dynamics contrast (KV-Diff from forte to piano). The result of the regression indicated that the model explained 60.7 % of the variance ($F(4, 31) = 14.5$, $p < .001$). It was found that the dynamics contrast from forte to piano (KV-Diff; β = -1.64, $p < .001$) was the only significant predictor of participants' judgments as teaching.

Piano. A multiple regression analysis was conducted to predict participants' judgments as teaching based on performance features of tempo (IOIs), articulation (KOT for piano parts), dynamics (KV for piano parts) and dynamics contrast (KV-Diff from piano to forte). The result of the regression indicated that the model explained 49.5 % of the variance ($F(4, 31) = 9.57$, $p < .001$). It was found that tempo (IOIs; β = 0.55, p = .02) and dynamics contrast from piano to forte (KV-Diff; β = 1.09, $p < .001$) were significant predictors of participants' judgments as teaching.

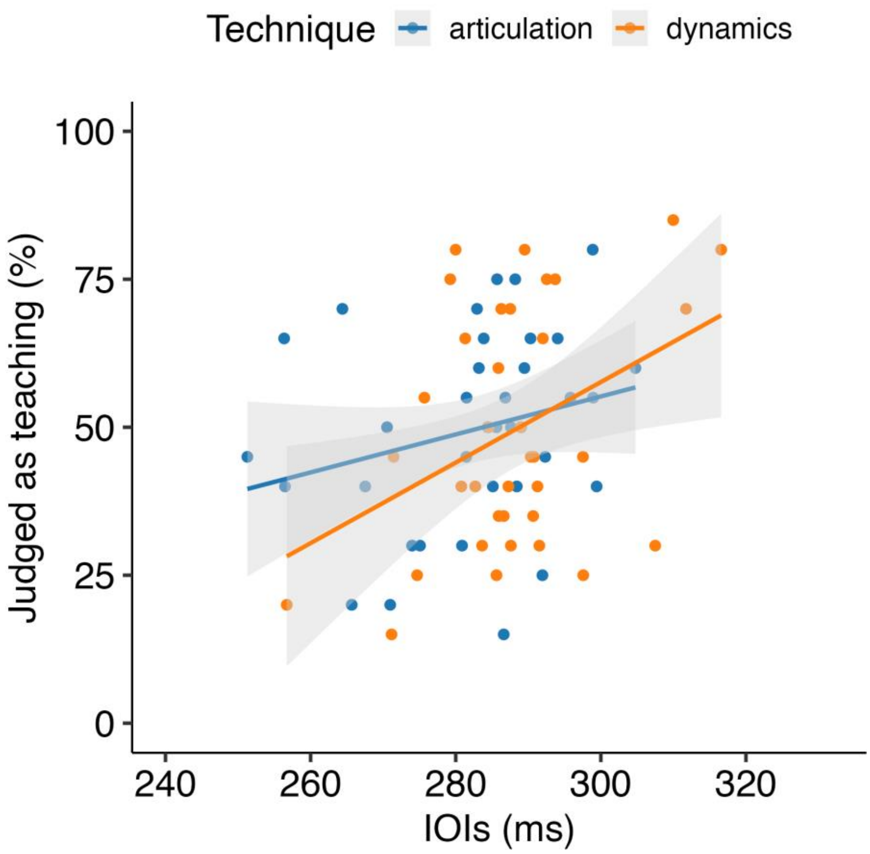

Figure 4.8. Experiment 2: Scatter plot showing the correlation between tempo feature (IOIs) and average participants' judgments as teaching for each recording. Therefore, each dot represents each stimulus. The grey bands display the 95 percent confidence intervals.

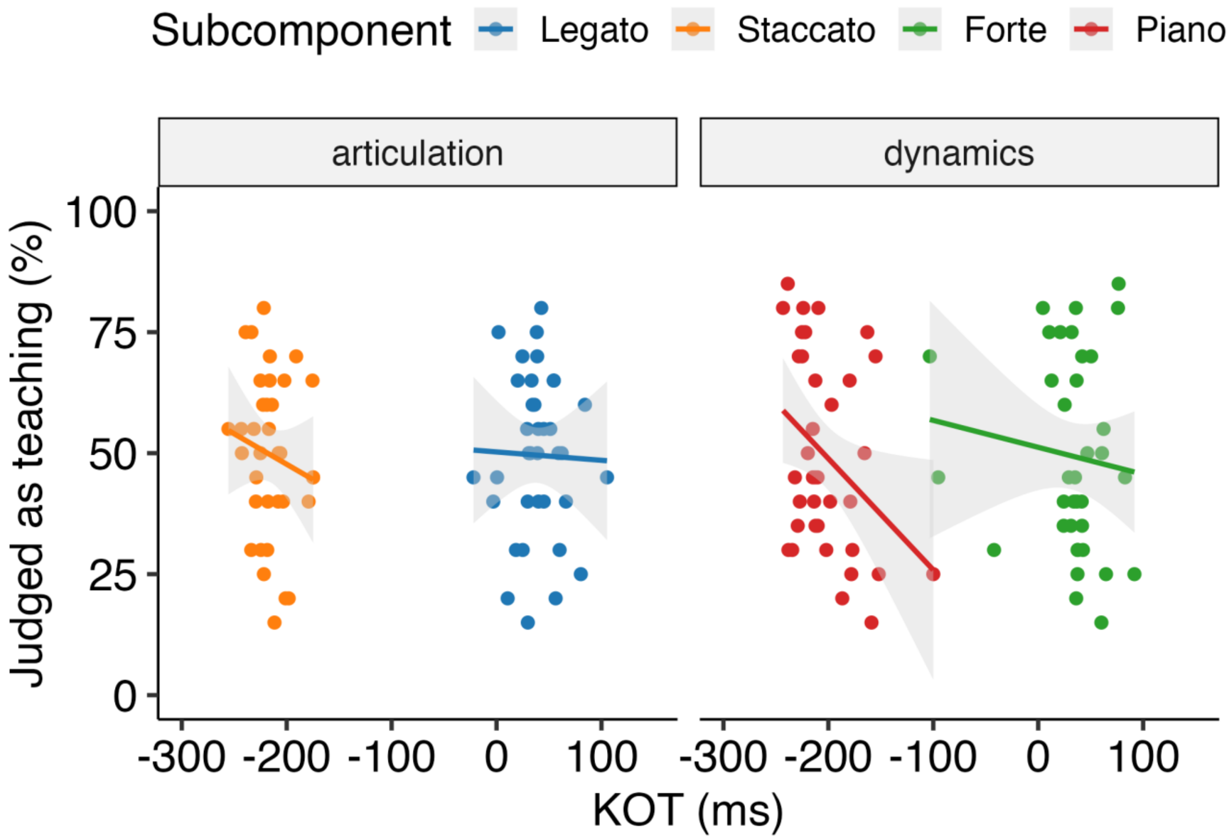

Figure 4.9. Experiment 2: Scatter plot showing the correlation between articulation features (KOT) and average participants' judgments as teaching for each recording. Therefore, each dot represents each stimulus. The grey bands display the 95 percent confidence intervals.

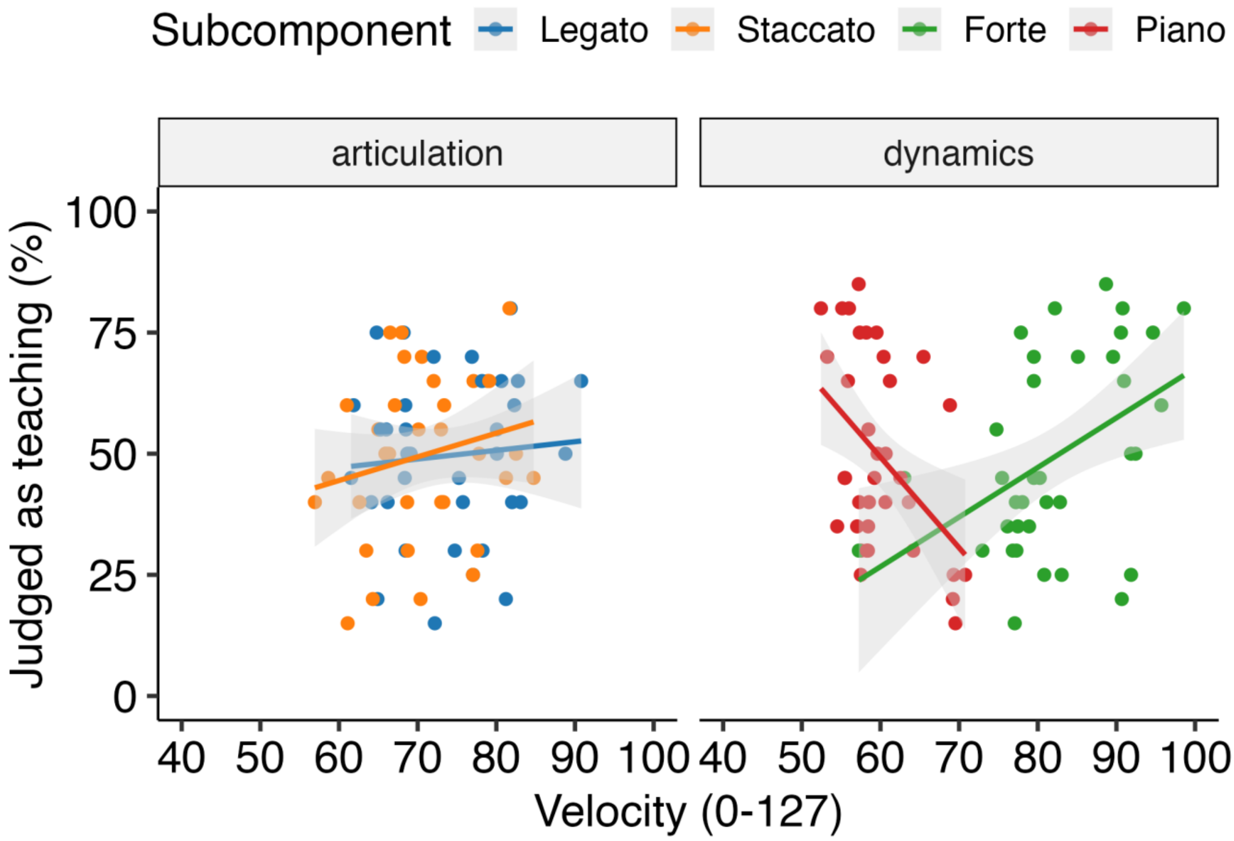

Figure 4.10. Experiment 2: Scatter plot showing the correlation between dynamics features (KV) and average participants' judgments as teaching for each recording. Therefore, each dot represents each stimulus. The grey bands display the 95 percent confidence intervals.

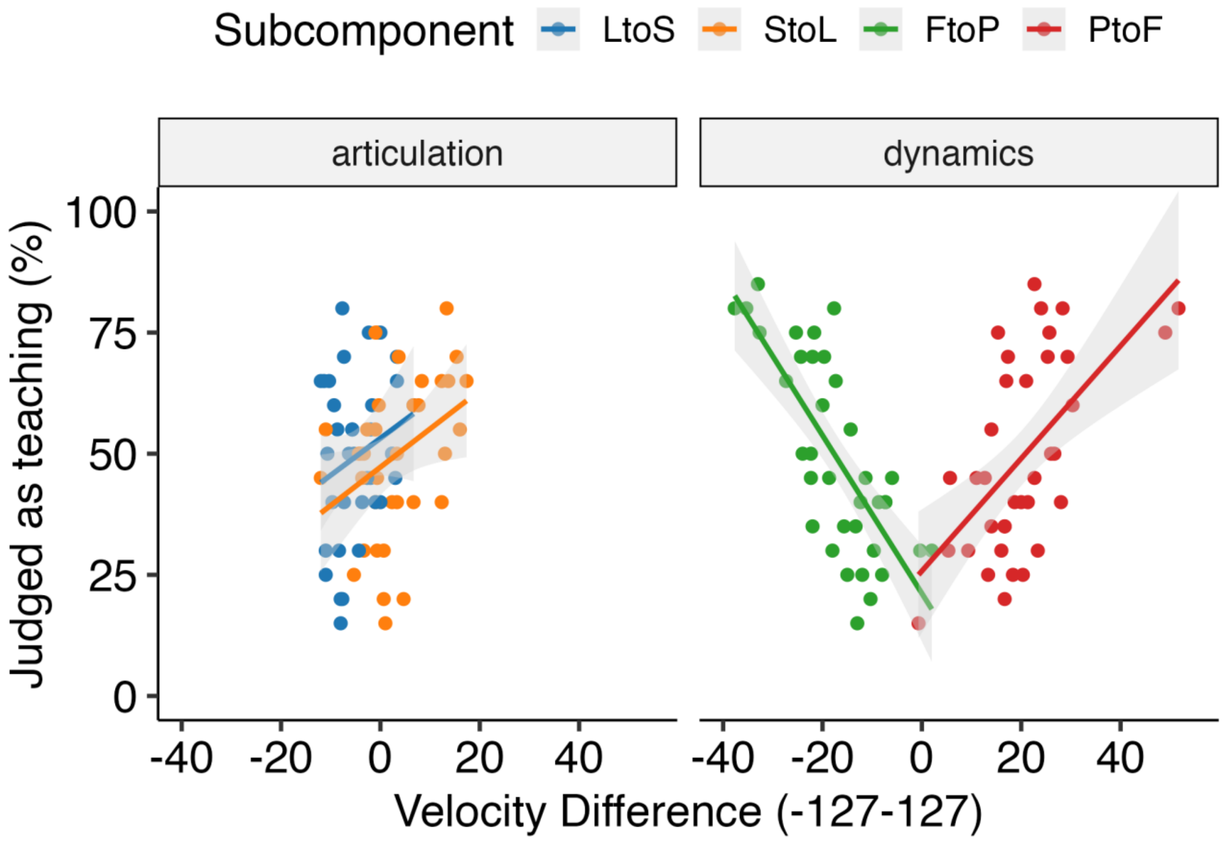

Figure 4.11. *Experiment 2: Scatter plot showing the correlation between dynamics contrast features (KV-Diff) and average participants' judgment as teaching for each stimulus. The grey bands display the 95 percent confidence intervals. LtoS; Legato to Staccato, StoL; Staccato to Legato, FtoP; Forte to Piano, PtoF; Piano to Forte.*

4.3.3 Discussion

The aim of Experiment 2 was to replicate the results of Experiment 1 and examine how musicians would infer teaching intentions when listening to a more complex piece of music. We found that articulation (KOT) did not seem to contribute to participants' judgments as teaching. Also, tempo (IOIs) was related to participants' judgments as teaching for dynamics recordings only. It seems that variations in these two parameters were considered to be in the space of possible interpretations of the music and were not used as general cues to infer teaching intentions.

However, dynamics (KV) seems to continue to be used as a cue to teaching intentions despite the increased musical complexity. Particularly performances with larger contrasts between forte and piano seem to be considered as teaching. Moreover, performances with both exaggerated forte and piano were considered to be for teaching. Thus, loudness (dynamics) might be used as a reliable cue to infer teaching intentions regardless of the complexity of a musical piece.

4.4 General Discussion

The present study investigated whether and how musicians infer pedagogical intentions by listening to piano recordings. In both Experiment 1 and 2, participants were able to choose correct answers accurately more than chance (50%). However, the results showed that the accuracy was relatively low (52.7 % in Experiment 1; 52.8 % in Experiment 2). This might be because the stimuli were randomly sampled from the teaching and performing condition in Tominaga et al. (2022). There was great variability in the selected recordings from multiple pianists and some of them do not seem to have exhibited the characteristics of teaching performance (e.g., slower demonstration, exaggeration). A possible reason is that some participants had difficulties in situating themselves in a teaching context and therefore did not produce cues.

The reason why we randomly selected the stimuli from each condition (i.e., teaching or performing) was because all the characteristics of teaching performance (e.g., IOIs, KOT, KV) were intercorrelated and it was difficult to choose stimuli based on one single criterion. Future research should examine if musicians answer more accurately with more experimentally controlled stimuli. Another possibility for the low accuracy could be that it was difficult even for musicians to find which cues were produced for teaching when listening to one piece of music with several phrases. Future research should examine whether accuracy would be improved when listening to phrases compared to whole pieces.

To examine which features of piano performance make musicians infer pedagogical intentions, we performed correlation and multiple regression analysis. Across the two experiments, it was found that loudness, particularly larger contrasts between forte and piano, strongly contributed to participants' judgments when listening to dynamics recordings. On the other hand, unlike what we predicted, slower performance was only considered to be for teaching for both techniques in Experiment 1 where the stimuli consisted of simple musical scales. Also, we could not find that performances with exaggerated articulation (i.e., longer legato and shorter staccato) were more likely to be judged as teaching in Experiment 2. This indicates that some characteristics of performance are not necessarily used or reliable to infer performers' intentions when listening to musically complex pieces. Certain modulations may have been too subtle to be noticed by the participants. This would be in line with Gabrielsson and Lindström (2010) who reported that tempo and loudness reign over other musical factors and seem to be easier for listeners to judge. Another possibility is that a wide range of modulations is considered as expressive in more complex pieces. This raises the possibility that it is easier to learn and teach musical expressive techniques when musical complexity is low.

The piece used in Experiment 2 was from actual sonatinas. Although sonatina pieces are generally aimed to be simple and easy for beginner pianists to play, they add a multitude of musical properties to a performance that is not present in musical scales. Not only do performers produce their communicative intentions such as emotion, but composers can express their feelings by handling various musical notations such as melody, rhythm and harmony (Gabrielsson & Lindström, 2010). In our study, most of the staccato parts of the piece used in Experiment 2 (*Fig 4.7* (2)) were comprised of the same repetitive notes. This implied that they could be only played in a staccato style. Dynamics may be more independent of the structure of the piece compared to articulation and relatively easier to perceive (Nakamura, 1987). Therefore, participants might have had difficulties in recognising teaching intentions from articulation in Experiment 2. This would imply that the ease or difficulty with which teaching intentions can be derived varies across different pieces depending on a variety of musical properties.

In the current study, we exclusively recruited musicians to explore our research questions. The reason why we recruited musicians only was because the concepts of articulation and dynamics seem to be difficult for those who don’t play an instrument to understand in the current experimental settings. It would be important to investigate how complete novices perceive and infer pedagogical intentions differently from musicians, who already have some experience in playing music.

Data availability statement

All data is available at https://osf.io/f6nr2/.

Chapter 5. Discovering musical expression together with teachers: A qualitative approach

5.1 Introduction

In this section, I will explore qualitative aspects of expressive performance in music based on my interview research conducted between October and December 2021 in collaboration with Prof. Martin Clayton at Durham University in the United Kingdom, as well as my own experience as a hobby pianist. The aim and motivation of my interview research at Durham was to complement and extend my experimental studies of musical skill transmission by using qualitative methods to better understand teachers' ideas about what effective teaching is and learners' ideas about what constitutes successful learning. In particular, I was interested in teachers' subjective views about (teaching) musical expression and the use of sound/action exaggeration in their lessons.

I started learning the piano at the age of five in Japan and have been taught by three teachers in my life. The first teacher taught me up to the age of 12 and the second teacher taught me between the age of 12 and 15. After I entered high school, I stopped playing the piano. I resumed private lessons after I started my PhD in Hungary with the third teacher when I was 28 years old and continued until the COVID-19 pandemic happened. Although I have never learnt about musical expression as deeply as music students do in a conservatoire, I realised how each of my teachers taught expressive performance was unique and different from the others. For example, my first teacher was concerned about specific sound properties and quality of music, which were mostly related to physical techniques. My second teacher liked to use verbal metaphors and imagery to facilitate my imagination and interpretation of music. My third teacher was concerned about an effective fingering to use and spent most of our lesson time finding the best fingering for me. He also advised me to improve my expressive skills by *not* playing the piano (e.g., reading books, travelling). It seems to me that there are several layers in expressive performance itself as well as students' learning stages. My teachers prioritised particular aspects of expressive performance based on what they thought was important or necessary for me to learn.

Based on my personal experiences with my teachers, I developed five themes about expressive performance and conducted interviews with music teachers and students. First, I started asking about the interviewees' learning history to see how they developed their own views of music (I: learning history). Second, I asked when and how they started teaching students (II: teaching experience). Third, I asked how they teach musical expression to find what they think is important to teach (III: teaching expression). Fourth, I was interested in how teachers use sound or action exaggerations in their lessons (IV: the use of sound/action exaggeration). Particularly, I wondered if exaggerations are used for pedagogical signalling as I discussed mainly in Chapter 1 and 5. Lastly, I briefly asked about cultural differences in terms of teaching styles (V: cultural difference).

5.2 Methods

Participants Six music teachers and two music students participated in semi-structured interviews. Three teachers were pianists, two teachers were Hindustani raag singers and one teacher was a cellist. One student was a Hindustani raag learner and the other student was a cello learner. All participants gave their informed consent before the interview started. We recruited participants by sending an email to music teachers who were registered on the list at the Department of Music at Durham University in the United Kingdom. Each interview was conducted either online or in person. The study was approved by the Durham University Music Department Ethics Committee.

Material and data analysis Our qualitative study was carried out using a semi-structured interview method (DiCicco-Bloom & Crabtree, 2006; Kallio, Pietilä, Johnson, & Kangasniemi, 2016). We developed the predetermined questions to facilitate a dialogue with interviewees and there were five topics in the questions (I: learning history, II: teaching experience, III: teaching expression, IV: the use of sound/action exaggeration and V: cultural differences; see *Supplementary Material*). Participants' verbal responses were recorded via Zoom (Zoom Video Communications) for online interviews and a voice recorder for in-person interviews. The audio data were transcribed for data analysis. Here, I would like to share some findings of my subjective observations from the transcribed data, particularly focusing on the topics of learning/teaching expression (I, III) and the use of sound/action exaggeration (IV).

5.3 Results and Discussion

I identified four topics from my interview data to complement and extend the findings of our experimental studies.

5.3.1 Teachers' tradition and overimitation

One of the main themes that came to mind from my interviews was that teachers tended to associate how they acquired expressivity with the group they belonged to. For example, one interviewee (a piano teacher) explained how she developed artistic performance by describing what the traditional piano school (i.e., Russian school) taught her in terms of not only technical but also creative aspects of performance.

> "One of the best things of Russian school particularly was that it's a perfect combination between working disciplines and work ethics, but at the same time, creative way of thinking. That's where art happens . . . A specific way of touching

keys. Because of the specific way of seeing music and representing it, hands are only outside of that. Inside, there is a huge work . . . Art is also about performing. To get interesting ideas of how to perform differently. Perform the tradition but at the same time . . . bring something new into that. This is the Russian school thing . . . I would definitely say that students are something that I made myself. I made their hands, I mean pianistic hands. I gave them the pianistic habits."

National piano schools were established in Europe, which refers to specific performance practices and pedagogies represented by the Russian, French and German schools (Lourenço, 2010). National piano schools are described by national characteristics (e.g., culture, historical and political circumstances), interpretations through traditions (e.g., techniques, sound and aesthetics) and influential pedagogues and/or performers (Wisniewski, 2016). For instance, the Russian school is described as being very passionate (aesthetics) and disciplined (personality) whereas the German school is described as being structured (aesthetics) and not showy (personality) according to Wisniewski (2016)'s questionnaire study (Chapter 5). Lourenço (2010) analysed three recordings of the same piece (Beethoven's Sonata op. 57, Appassionata) played by three historical pianists (Vladimir Sofronitzky, Robert Casadesus, Edwin Fischer) who were supposed to represent each national piano school (Russian, French, German). The results demonstrated that the performances of the three pianists have distinct characteristics with regard to dynamics, tempo and polyphonic textures.

As Lourenço (2010) and Wisniewski (2016) demonstrated, national piano schools seem to have developed particular characteristics of expressive performances, which have been transmitted over generations. It can be considered that teachers of each school put an emphasis on certain aspects of performance that are considered to be important in each culture. As the interviewee above mentioned, it is not expected to copy what teachers have said or previous pianists have done. Students are expected to bring something "new" from their own interpretations. However, how students perceive, analyse and express music seems to be constrained by their teachers and their specific musical cultures to some extent.

Having a sense of belonging to one specific group may facilitate the learning of expressive performance, which consists of many domain-specific and culture-specific rules. In developmental psychology, it has been known that children tend to imitate causally irrelevant actions of a demonstrator to achieve an instrumental goal (i.e., overimitation; Lyons, Young, & Keil, 2007). The propensity to overimitate (copy faithfully) is supposed to contribute to transmitting cultural knowledge, skills and conventions (Legare & Nielsen, 2015; Shipton & Nielsen, 2015). The tendency to overimitate becomes stronger if a model who demonstrates actions belongs to the same social group (Altınok, Király, & Gergely, 2022; Gruber, Deschenaux, Frick, & Clément, 2019). Over and Carpenter (2013) argued that children have social motivations and feel social pressures to copy ingroup members to show their affiliation, which may lead them to reproduce faithfully. It would be interesting to investigate how expressive performance in music is copied from a teacher belonging to the same group and from a teacher belonging to a different group.

5.3.2 Exaggeration as an exploration tool

One of the concerns about the use of exaggeration in music is that students might copy exaggerated actions or sound without noticing teachers' demonstration was exaggerated to teach a specific aspect of a technique. However, this does not seem to be a problem in a direct teaching situation because teachers can correct students' performances or interpretations by giving feedback. One interviewee (a piano teacher) shared her experience with the use of exaggeration during her lesson.

> "I have come across a few times that I over-exaggerate to make a point in one lesson and then next, the following week, they (students) over-exaggerate and playing like with accents or like you know, it just goes beyond what I asked, which is fine, because that's what I asked to do, but then I ask them to do exactly the same, but just inside [of their mind] . . . I think anything with not just music but anything takes time to digest. And sometimes it's good to just experience the extreme version. And then, you balance out."

Experiencing the extreme version could be interpreted as experiencing "variability" in performance. This is in line with some findings from the domain of motor learning, showing that task-relevant motor variability contributes to faster learning (Dhawale, Smith, & Ölveczky, 2017; Wu, Miyamoto, Castro, Ölveczky, & Smith, 2014). Therefore, during a learning process where students are allowed to explore, exaggerated performance may make students try different possibilities in terms of action, sound and expression and work effectively for their learning processes.

5.3.3 The role of teachers

It seems that one of the important elements of teaching is to prepare an appropriate environment for students to explore and give some constraints if they exceeded boundaries. One interviewee (a piano teacher) talked about the role of teachers with an interesting metaphor.

> "The piano teacher should know the piano repertoire very well. You should also study your students, their mentality, their temperament . . . It was something like choosing the best-ever dress from your wardrobe to offer just enough [for] them to start . . . You need to find the exact colour, you need to find the exact style [for students]."

Another interviewee (a Hindustani raag singing teacher) also reported a similar thing by stating that “Ideally [like in a one-on-one session], teachers would take account of the student’s personality and think about what kind of melody would suit the student”. Although he mentioned that there are a couple of basic raags that most students start with, he explained that it is important for teachers to choose an appropriate raag depending on students’ abilities to facilitate learning. There is an enormous number of materials (songs, pieces) to study in music and it is difficult for students to know where to start to acquire skills effectively. Therefore, the ability to choose suitable materials for each student seems to be important for teachers.

One interviewee (a piano teacher) mentioned that she avoids giving direct instructions or suggestions on what to do and gives feedback when her students go beyond the limit of possible performance.

> “I only tried to give them hints and suggestions to provoke their own ideas and I talked about phrasing and how to change colour and touch and suggest where to take time, or you know, what we called meeting points should take place within the piece . . . For adult students, I encourage them to come up with their own ideas and obviously when it’s stylistically not correct or if it doesn’t make sense, as a music interpretation, then I tell them that it’s not right.”

This is in line with the findings of a study by Schiavio, van der Schyff, Philippe and Biasutti (2022), where music teachers inspired students to consider novel options while giving constraints to help students build creative activities. Legare and Nielsen (2015) discussed that children have difficulty finding new possibilities and solutions at the early stage of development and that innovation needs didactic pedagogy and scaffolding from others such as caregivers and peers. Therefore, innovative behaviour such as finding new ideas and options may not occur through random trial-and-error but requires exploration strategies, which are acquired through interacting with experts.

5.3.4 Never-ending learning

One of the common themes that I observed in my interviews was that teachers also considered themselves as learners and the learning process of music never ends as long as you keep performing music. One interviewee (a Hindustani raag singing teacher) described why they never finish learning by explaining the learning process of Hindustani classical music with an interesting analogy.

> "We often say like learning a raag is like getting introduced to a person. So when you are getting introduced to a person, at first, we start by knowing the person's name and then where the person comes from or what the person's interests are . . . Similarly, the more you get attached to a raag, the more you listen to it, the more [you] try to listen and produce it, [you] learn more about the raag as you would learn about a person . . . The learning trajectory would be quite different because when you are learning a raag as a 5-year-old, compared to when you are learning as a 15-year-old and then 25-year-old and later on when you are 40 . . . You're never done with anything in this system of music (Hindustani classical music) . . . You tend to come back to a [basic] raag many times in your learning life."

Another interviewee (a piano teacher) mentioned similarly that improving yourself is important as a professional musician.

> "I listen to different pianists playing the same piece to get different ideas . . . I enjoy the inspiration from the other instruments [in Chamber music], how they are phrasing it or how they express . . . I think it's important as a musician that you always want to

improve yourself. Always want to go higher than where you are . . . Sharing music and interacting with other musicians definitely advances your musicianship."

Creating new and unique ideas is vital to develop as a musician. Schiavio et al., (2022) discussed that the ability to see things from a different perspective is considered to be a valuable factor for music creativity. How people change their perspectives seems to rely on two factors: the way you develop as an individual and the way you interact with others. With regard to the first factor, you never stop growing as a human being and cannot help having a new perspective as you experience different things in your life. With regard to the second factor, interacting with others seems to give an opportunity for you to think about why you interpret music in a certain way by comparing it with other people. Teachers play a role as "others" by giving students opportunities to perceive, analyse and perform music and peers also start influencing each other. Therefore, even if you are performing music alone as a solo pianist, your musicality has been created with others throughout your musical life (Schiavio, Ryan, Moran, van der Schyff, & Gallagher, 2021).

5.4 Conclusion

The aim of this section was to share insights from my qualitative study and to extend our findings of three experimental studies by discussing additional aspects of expressive performance in music. I discussed how teachers influence students with regard to their learning processes and performance itself by examining four key observations from my interviews. First, I discussed how teachers developed their expressivity through the group that they belong to and implied that overimitation in a social context might contribute to acquiring expressive skills. Second, I examined how teachers used action/sound exaggerations in their lessons and found that expression was used not only for pedagogical signalling but also for letting students experience variability in expressive performance. Third, I discussed the role of teachers as guides while students are exploring their music. Lastly, I mentioned that learning never ends and that musicians' expression has been developed through interactions with other people such as experts and peers.

One of the limitations of this study is that my subjective observations cannot be generalised due to the nature of this qualitative work. I interpreted the results by considering them in the context of findings in cognitive science (e.g., developmental psychology, cognitive psychology), such as overimitation and motor variability in skill acquisition, in order to generate further ideas for new experiments. However, there are many interesting findings that I could not include here such as how experts teach differently to adult and child students and how they use different materials to enhance imagination (e.g., the use of visual art). To make this analysis more formal, it would be necessary to analyse data with thematic analysis (Braun & Clarke, 2012) by multiple people to ensure objectivity. Lastly, it was not possible to observe actual lessons because of the COVID-19 pandemic situation as well as the nature of my short-stay research, which made it difficult to have enough time to establish a deeper rapport with the teachers I interviewed. It would be interesting to compare what teachers described in the interviews with what they actually do.

Funding acknowledgement

The research supporting this section was fully sponsored by Central European University Foundation of Budapest (CEUBPF). The theses explained herein represent the ideas of the author, and do not necessarily reflect the views of CEUBPF.

5.5 Supplementary Material

I. Learning history

1. When did you start playing music/instruments?
2. Why/how did you start learning music/instruments?
3. Why did you choose those instruments?
4. Where did you receive formal training?
5. How many teachers did you have?
6. Who was the most impressive teacher for you?
7. When and how are you trained to be a teacher?

II. Teaching experience

8. Do you plan what/how to teach before each lesson?
9. What is your structure of lessons?
10. What do you think is important for students to learn (let's say top 3)?
11. How do you give feedback to your student? (e.g., gestures, verbal instructions, use of imagery, yes or no)
12. How do you evaluate your students' improvements? (e.g., compared to previous sessions, compared to your performance, compared to an ideal performance)
13. How do you select homework for your student?
14. When do you feel that you have been successful in teaching?
15. What kind of feedback is easier to understand for students?
16. Is there any difference between when teaching adults and when teaching children?
17. What do you expect students to learn?
18. What do you think is the ultimate goal of learning music?

III. Teaching expression

19. What do you do when teaching articulation/dynamics to students?
20. Which is easier for students to learn, articulation or dynamics?
21. Is singing useful to teach musical expression?
22. Do you use metaphors to describe musical expression?

IV. The use of sound/action exaggeration

23. Do you exaggerate your performance for teaching purposes? (e.g., playing forte more than it should be)
24. Do you play slower for students?
25. Do you deliberately try to use pointing or have eye contact with students?
26. Do you decompose a piece of music for students? If so, do you have any rules? (e.g., measure by measure, bar by bar, theme by theme)
27. Do you reconstruct the whole piece for students or explain the whole structure of the piece to students?

V. Cultural difference

28. Have you faced or noticed any cultural differences between teaching traditions in the UK conservatories and those in your own country?
29. Do you think your cultural background affects your teaching style?

Chapter 6. General Discussion

This dissertation aimed to investigate the role of teaching in skill transmission. Particularly, I focused on pedagogical sound modulations when teaching expressive techniques in music. In Chapter 1, I reviewed important elements of teaching and proposed three experimental projects to investigate how experts produce pedagogical cues for novices to highlight relevant information for skill acquisition (Chapter 2), how experts adapt their pedagogical signalling depending on the skills of novices (Chapter 3) and how these pedagogical signals are detected and potentially used by novices (Chapter 4). In addition to my experimental work, I explored creative aspects of musical expression by conducting interviews with music teachers and students (Chapter 5). In this chapter, I will summarise our results and discuss the implications of our findings and the limitations of our approach, followed by future directions.

In Chapter 2, I investigated whether and how expert pianists modulate their performance when they have the intention to teach. In two experiments, participants were asked to perform one piece of music with notated expressive techniques of either articulation (the smoothness of sound) or dynamics (the loudness of sound). We compared participants' performances when they intended to teach a designated expressive technique to students and when they tried to perform it to an audience. Our findings showed that expert pianists exaggerated relevant aspects of the techniques to be taught. In particular, participants tended to produce shorter staccato when teaching articulation and larger contrasts between forte and piano when teaching dynamics. These findings suggest that expert pianists systematically modulate their sound to highlight relevant information to novice learners.

In Chapter 3, I examined whether and how expert pianists adapt their performance depending on the skills of novice learners. In order to maximise experimental control, we created artificial recordings of novice learners. We manipulated the extent to which articulation and dynamics were implemented in the recordings. The recordings with both techniques implemented were considered to resemble good student performances whereas those without any implemented techniques were considered to resemble poor student performances. In the experiment, expert pianists were asked to listen to each recording and then perform the same piece for each student for teaching purposes. Our findings showed that expert pianists prioritised the teaching of articulation over that of dynamics. As predicted, they exaggerated their performance with regard to articulation when students did not produce articulation correctly. However, they exaggerated their performance with regard to dynamics only when students correctly produced articulation. These results may indicate that one technique might be more important than the other, at least for the particular musical piece that we used. The findings could also imply that experts focus their teaching on one aspect at a time rather than teaching multiple techniques together, which could be investigated in future studies.

In Chapter 4, I investigated whether musician listeners (i.e., in the role of potential learners) are able to discriminate between recordings with and without pedagogical intentions by listening. Here, I examined which performance features contributed the most to judging that a recording was produced for teaching. Our results showed that musician listeners seemed to think that slower performances were produced for teaching in general. Moreover, they judged that performances with exaggerations were for teaching. However, listeners did not infer teaching intentions from performances with exaggerated articulation when the complexity of a piece was high. This might be because listeners considered that exaggerated articulation reflected the interpretation of the piece rather than sending pedagogical signals. These findings suggest that listeners use particular performance features to infer pedagogical intentions from the recordings, which indicates that pedagogical cues produced by experts might be picked up by novice learners.

In Chapter 5, I focused on creative aspects of expressive performance by conducting interviews with professional music teachers and students. I identified four key observations from my interviews. I found that action/sound exaggeration can be used not only for pedagogical signalling but also to allow students to explore their expressivity. Also, I discussed that the role of teachers is to give feedback when students go beyond possible interpretations, rather than to give direct instructions and suggestions on what to do.

6.1 Ostensive communication in skill transmission

Our findings are consistent with the theory of natural pedagogy, which proposes that ostensive communication has developed to transmit generic knowledge and skills between individuals (Csibra & Gergely, 2006, 2009, 2011). Ostensive communication is achieved when a communicator produces a stimulus (e.g., speech, action) for an addressee with informative and communicative intentions to deliver relevant information and to convey that the information is relevant to the addressee specifically. The addressee needs to infer the communicator's intentions and extract the information from the stimulus (Sperber & Wilson, 1995). In my studies, I consider sensorimotor communication as one form of ostensive communication in the sense that it plays a role in signalling information (i.e., knowledge and skills to be acquired) with communicative intentions (i.e., expressed through exaggerated actions) from experts (i.e., more knowledgeable partners) to novices (i.e., less knowledgeable partners) for skill transmission (e.g., Vesper, Morisseau, Knoblich, & Sperber, 2021). Of course, novices may be able to acquire knowledge and skills just by observing others relying on general learning mechanisms such as statistical learning (Saffran & Kirkham 2018). However, it is often the case that actions we perform have hierarchically complex structures and may contain cognitively opaque cultural knowledge (Botvinick, 2008; Heyes, 2017). Experts' ostensive communication is supposed to be useful for novices to acquire knowledge and skills efficiently because experts indicate what to acquire by highlighting relevant information with ostensive cues such as eye contact, gestures and contingent turn-taking (Csibra & Gergely, 2006). In Chapter 2 and 3, we found that expert pianists highlighted relevant aspects of performance (e.g., to produce shorter staccato) according to an instructed goal (e.g., to teach articulation to students) even when they were only allowed to communicate through sound. In naturalistic teaching settings, expert pianists likely produce further ostensive cues such as eye contact, pointing and gestures (Simones, 2019; Simones et al., 2015), which could also help novices' skill acquisition.

Natural pedagogy proposes that novices need to make inferences about experts' intentions from their ostensive demonstrations to extract relevant information efficiently. In general, novices have difficulty in identifying relevant knowledge and skills from observation because there is an infinite number of possible goals to infer from novices' perspectives (Csibra & Gergely, 2006). Ostensive communication is supposed to provide constraints on novices' cognitive mechanisms to narrow down the possible goals of observed actions. Indeed, it has been found that novices seem to process information differently when it is demonstrated with and without ostensive cues (Egyed, Király, & Gergely 2013; Király et al., 2013; Okumura et al. 2020; Yoon2008). For example, Király et al. (2013) investigated how infants imitate actions differently from communicative and non-communicative demonstrations. Infants observed a model pressing a lamp in order to activate it (the lamp was lit while being pressed) in two conditions. In the Hands Free condition, the model pushed the lamp by her head while her hands were free. In the Hands Occupied condition, the model pushed the lamp by her head while her hands were covered by a blanket (i.e., her hands were occupied). In Experiment 1, the model ostensively communicated with the infants by calling their names and having eye contact. The results showed that infants imitated the head action more in the Hands Free condition than in the Hands Occupied condition[2]. However, this difference was not observed when infants incidentally observed the model producing the head action (e.g., in a non-communicative context; Experiment 2). In their study, the model always produced the head action to activate the lamp but the exact action was imitated only in a communicative context. Therefore, what people learn (e.g., recognise, memorise or imitate) from demonstrations may depend on how teachers demonstrate. In Chapter 4, we found that musician listeners inferred the pedagogical intentions of performers by listening to recordings. The findings indicated that ostensive cues (e.g., exaggerated dynamics) produced by experts were interpreted as reflecting teaching intentions by novices. Future research should investigate whether and how novices can learn (e.g., memorise or imitate) differently from pedagogical (e.g., teachers' demonstrations) and non-pedagogical demonstrations (e.g., ideal performances by professional musicians).

Generalisability is also an important factor in natural pedagogy, which I have not examined in this dissertation. Natural pedagogy is supposed to aid with the transmission of generalisable knowledge and skills, which are not restricted to particular episodic facts but convey useful information over generations (Csibra & Gergely, 2009). Although musical expressions might be seen as arbitrary and unique to each musical piece and each performer, there are some similarities between performers in terms of expressions (Repp, 1992, 1997). In the case of expressive performance in music, techniques such as articulation and dynamics

[2] Gergely, Bekkering and Király (2002) originally found that infants tended to produce the head action more in the Hands Free condition than in the Hands Occupied condition. They explained that infants inferred that the model used her head because her hands were not available in the Hands Occupied condition and she needed to use her head instead to activate the lamp. Therefore, infants used their hands instead of their head when they were asked to play with the lamp by themselves. On the other hand, in the Hands Free condition, infants inferred that it was important to activate the lamp with a head, not with hands because the model used the head even when her hands were available. Therefore, infants imitated the head action more in the Hands Free condition than in the Hands Occupied condition.

are generalisable skills because musicians are required to employ these expressive tools to demonstrate their emotions and interpretations of music (Bernays & Traube, 2014). It would be interesting to study what characteristics of performance are likely to be taught, learnt and transmitted over generations.

6.2 Interactivity between experts and novices

Our current approach started with a minimal teaching scenario where experts were asked to play for imaginary students (Chapter 2). In Chapter 3, we enriched this by having experts listen to recordings of students, but there was still no reciprocal interaction between experts and novices. In Chapter 4, we turned things around by asking listeners to identify teaching intentions, but there was again no interaction. This approach allowed us to investigate our research questions focusing on sensorimotor communication without considering other social factors.

One of the limitations of our approach is that the information flow between experts and novices was unidirectional whereas, in naturalistic settings, experts and novices interact with each other reciprocally. In order to address reciprocity in communication between experts and novices, our studies could be extended to investigate the effects of turn-taking between experts and novices. For example, Pan and colleagues (2018) examined how the frequency of turn-taking learning between instructors and participants affected participants' performance after a learning session. In the experiment, participants were asked to learn a song that consisted of four phrases. The degree of turn-taking was manipulated so that in one condition, an instructor demonstrated one phrase and a participant imitated it (i.e., turn-taking per phrase) whereas, in the other condition, an instructor demonstrated an entire song and a participant imitated it (i.e., turn-taking per song). Therefore, the first condition (Part Learning (PL) condition) had more interactions than the second condition (Whole Learning (WL) condition) to demonstrate one song. They found that participants' performance was significantly better in the PL condition than in the WL condition, suggesting that more turn-taking with experts is beneficial for skill acquisition.

Even though turn-taking is a prototypical way of engaging in dyadic interactions (Brownell, 2011; Gratier et al., 2015; Levinson, 2016), synchronous performance also plays a role in the interpersonal coordination between experts and novices (Phillips-Silver & Keller 2012). For example, piano teachers sometimes sing or count beats while students are performing in order to teach the sound quality of a particular expression or to deliver information related to tempo (Speer, 1994). Performing together with experts may be effective as novices are able to entrain to the beats produced by experts and acquire particular techniques requiring precise temporal coordination such as expressive timing (Knoblich, Butterfill, & Sebanz, 2011). In naturalistic teaching situations, the two types of joint action (turn-taking vs. synchronous playing) are intermixed. Although many studies on skill transmission seem to focus on turn-taking scenarios where experts demonstrate first and novices imitate, future studies should investigate how performing together may contribute to facilitating novices' skill acquisition.

Another limitation of our approach is that how experts and novices could communicate was strictly constrained, namely only through sound. In actual teaching settings, expert pianists communicate not only through sound demonstration but use verbal and non-verbal communication such as gestures (Simones, 2019; Simones et al., 2015). For instance, Li and Timmers (2020) reported that pianists understand piano timbre as an embodied concept, which is shaped by not only sound properties but also visual, haptic and kinematic experiences. Therefore, future studies should investigate how the use of communication channels other than sound demonstration facilitates novices' learning and enhances their performance.

6.3 Identifying the goal of expressive performance

Natural pedagogy assumes that transmitted knowledge or skills have the function to achieve a certain goal (Csibra & Gergely, 2006). Teaching is effective because experts know how to obtain a particular goal and modulate their demonstration so that novices are able to understand the goal efficiently. In this dissertation, I investigated the teaching of fundamental expressive techniques (articulation and dynamics) that are well-defined conceptually and technically. For example, expert pianists know what articulation means and can demonstrate how to hold keys, move their fingers and use their wrist to produce smooth or separated sound. It has also been discussed that successful expressive performance is achieved by an explicit plan or goal with the idea of concrete sound properties (Johnson, 2000; Woody, 2003). I speculate that the technical component of expressive performance, which is related to particular motor skills (Sloboda, 2000), can be generalisable and hence transmitted to other individuals. Sensorimotor communication is a special type of ostensive situation in which information and communicative intentions are embedded in the same instrumental action. Other ostensive cues such as eye contact and pointing may guide novices' attention to relevant information but the cues themselves do not contain any practical information to execute the action to be acquired. Sensorimotor communication is considered to help novices develop a common representation or goal of perceived actions that experts produce and performed actions that they imitate by participating in joint action with experts (McEllin, Knoblich, & Sebanz, 2018).

On the other hand, the ultimate goal of teaching expressivity is that every musician develops their own expression and interpretation of music (Meissner, 2021). The creative component of expressive performance seems to be acquired through a musician's own experience. Schiavio et al. (2022) found that music teachers try to foster creative attitudes in their students and let the students explore to find their own perspectives to facilitate creative musicality. Experts do not share their perspectives or goals with novices but rather, they seem to create a new perspective or goal together with novices by actively participating in the learning process. In Chapter 1, I argued that experts deconstruct skills for novices, and this deconstruction process may allow experts to explore the learning process again through the bodies of novices. In such a situation, the role of experts may not be fixed and novices themselves may take charge of their own learning (Schiavio, Biasutti, van der Schyff, & Parncutt, 2020). This is in line with my qualitative observations, where music teachers often said that they would never stop learning even after they had already established professional careers as teachers (*Chapter 5*). Also when teaching or exploring the creative aspects of expressive performance, action exaggeration seems to be used for different purposes. In a qualitative study of teaching, I found that some musical teachers said that they use exaggeration not only because students do not notice what to learn, but also because they want their students to experience some extremes or variability during the learning process (*Chapter 5*). Therefore, exaggeration in the context of teaching musical expressions seems to have multiple different functions.

The traditional teacher-directed model in conventional Western education has been criticised because it may hinder the creative and exploratory learning processes of novices (Borgo, 2007; Hickey, 2009). When it comes to teaching free improvisation, the focus of teaching seems to be on facilitating creative thinking and growth rather than giving direct instructions (Hickey, 2015). Also, learning from experts is not the only way to acquire knowledge and skills. For example in popular music, it is common to learn music through self-directed and self-assessed activities and develop musicianship through interacting with peers and engaging in group activities (Lebler, 2008). Although learning on your own and from peers is an important aspect and process of skill acquisition, it is also true that the traditional model has contributed to supporting cultural transmission and passing on knowledge and skills through experts' demonstration and novices' imitation. In real-world learning settings, different forms of teaching are intermixed and used. It would be fruitful to further explore how different teaching methods are employed and effective depending on various factors such as novices' age and skill levels (Bonastre & Timmers, 2021). For instance, the traditional model might be useful for beginners who are going to acquire basic skills to play an instrument whereas free exploratory learning would be beneficial for intermediate or advanced students who are already equipped with a basic skill set.

Music is one of the ubiquitous inventions in human culture. From early development, infants are sensitive to musical features of sound and seem to be motivated to engage in musical interactions with adults (Trehub, 2003; Phillips-Silver & Keller, 2012). From childhood to adulthood, humans expand their music activities from listening to performances to making music together with others. Although forms of music vary widely across cultures and societies, music is closely tied to our social life such as infant care, healing, dance and love (Mehr et al., 2019). It can be speculated that our domain-general cognitive mechanisms may be employed to perceive, produce and interact with music while how people realise the music in society is unique and embedded in a cultural context, involving domain-specific rules. I believe that this dissertation contributed to the study of musical skill transmission by investigating how domain-general cognitive mechanisms such as ostensive communication and sensorimotor communication are involved. Our findings reveal the importance of action-based communication between experts and novices, which can take place in the absence of verbal communication.

References

Aglioti, S. M., Cesari, P., Romani, M., & Urgesi, C. (2008). Action anticipation and motor resonance in elite basketball players. *Nature Neuroscience*, *11*(9), 1109–1116. https://doi.org/10.1038/nn.2182

Akkermans, J., Schapiro, R., Müllensiefen, D., Jakubowski, K., Shanahan, D., Baker, D., … Frieler, K. (2019). Decoding emotions in expressive music performances: A multi-lab replication and extension study. *Cognition and Emotion*, *33*(6), 1099–1118. https://doi.org/10.1080/02699931.2018.1541312

Altınok, N., Király, I., & Gergely, G. (2022). The Propensity to Learn Shared Cultural Knowledge from Social Group Members: Selective Imitation in 18-month-olds. *Journal of Cognition and Development*, *23*(2), 273–288. https://doi.org/10.1080/15248372.2021.1966013

Bernays, M., & Traube, C. (2014). Investigating pianists' individuality in the performance of five timbral nuances through patterns of articulation, touch, dynamics, and pedaling. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00157

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, *2*(8), 561–567. https://doi.org/10.1038/35086023

Bonastre, C., & Timmers, R. (2021). Comparison of beliefs about teaching and learning of emotional expression in music performance between Spanish and English HE students of music. *Psychology of Music*, *49*(1), 108-123. https://doi.org/10.1177/0305735619842366

Borgo, D. (2007). Free jazz in the classroom: An ecological approach to music education. *Jazz perspectives*, *1*(1), 61-88. https://doi.org/10.1080/17494060601061030

Botvinick, M. M. (2008). Hierarchical models of behavior and prefrontal function. *Trends in Cognitive Sciences*, *12*(5), 201–208. https://doi.org/10.1016/j.tics.2008.02.009

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for ‘motionese’: Modifications in mothers’ infant-directed action. *Developmental Science*, *5*(1), 72–83. https://doi.org/10.1111/1467-7687.00211

Brand, R. J., & Shallcross, W. L. (2008). Infants prefer motionese to adult-directed action. *Developmental Science*, *11*(6), 853–861. https://doi.org/10.1111/j.1467-7687.2008.00734.x

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association. https://doi.org/10.1037/13620-004

Brand, R. J., Baldwin, D. A., & Ashburn, L. A. (2002). Evidence for ‘motionese’: Modifications in mothers’ infant-directed action. *Developmental Science*, *5*(1), 72–83. https://doi.org/10.1111/1467-7687.00211

Bresin, R., & Battel, G. U. (2000). Articulation Strategies in Expressive Piano Performance Analysis of Legato, Staccato, and Repeated Notes in Performances of the Andante Movement of Mozart’s Sonata in G Major (K 545). *Journal of New Music Research*, *29*(3), 211–224. https://doi.org/10.1076/jnmr.29.3.211.3092

Brownell, C. A. (2011). Early Developments in Joint Action. *Review of Philosophy and Psychology*, *2*(2), 193–211. https://doi.org/10.1007/s13164-011-0056-1

Byrne, R. W. (2003). Imitation as behaviour parsing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 529–536. https://doi.org/10.1098/rstb.2002.1219

Byrne, R. W., & Rapaport, L. G. (2011). What are we learning from teaching? *Animal Behaviour*, *82*(5), 1207–1211. https://doi.org/10.1016/j.anbehav.2011.08.018

Campisi, E., & Özyürek, A. (2013). Iconicity as a communicative strategy: Recipient design in multimodal demonstrations for adults and children. *Journal of Pragmatics*, *47*(1), 14–27. https://doi.org/10.1016/j.pragma.2012.12.007

Cancino-Chacón, C. E., Grachten, M., Goebl, W., & Widmer, G. (2018). Computational models of expressive music performance: A comprehensive and critical review. *Frontiers in Digital Humanities*, *5*, 25. https://doi.org/10.3389/fdigh.2018.00025

Candidi, M., Curioni, A., Donnarumma, F., Sacheli, L. M., & Pezzulo, G. (2015). Interactional leader–follower sensorimotor communication strategies during repetitive joint actions. *Journal of The Royal Society Interface*, *12*(110), 20150644. https://doi.org/10.1098/rsif.2015.0644

Cavallo, A., Koul, A., Ansuini, C., Capozzi, F., & Becchio, C. (2016). Decoding intentions from movement kinematics. *Scientific Reports*, *6*(1), 37036. https://doi.org/10.1038/srep37036

Colprit, E. J. (2000). Observation and Analysis of Suzuki String Teaching. *Journal of Research in Music Education*, *48*(3), 206–221. https://doi.org/10.2307/3345394

Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. *Processes of change in brain and cognitive development. Attention and performance XXI*, *21*, 249-274.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, *13*(4), 148–153. https://doi.org/10.1016/j.tics.2009.01.005

Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1149–1157. https://doi.org/10.1098/rstb.2010.0319

De Felice, S., Hamilton, A. F. de C., Ponari, M., & Vigliocco, G. (2023). Learning from others is good, with others is better: The role of social interaction in human acquisition of new knowledge. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *378*(1870), 20210357. https://doi.org/10.1098/rstb.2021.0357

Dhawale, A. K., Smith, M. A., & Ölveczky, B. P. (2017). The Role of Variability in Motor Learning. *Annual Review of Neuroscience*, *40*(1), 479–498. https://doi.org/10.1146/annurev-neuro-072116-031548

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education*, *40*(4), 314–321. https://doi.org/10.1111/j.1365-2929.2006.02418.x

Egyed, K., Király, I., & Gergely, G. (2013). Communicating Shared Knowledge in Infancy. *Psychological Science*, *24*(7), 1348–1353. https://doi.org/10.1177/0956797612471952

Ericsson, K. A. (2008). Deliberate Practice and Acquisition of Expert Performance: A General Overview. *Academic Emergency Medicine*, *15*(11), 988–994. https://doi.org/10.1111/j.1553-2712.2008.00227.x

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, *8*(2), 181–195. https://doi.org/10.1016/S0163-6383(85)80005-9

Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, *10*(3), 279–293. https://doi.org/10.1016/0163-6383(87)90017-8

Fukuyama, H., Qin, S., Kanakogi, Y., Nagai, Y., Asada, M., & Myowa-Yamakoshi, M. (2015). Infant’s action skill dynamically modulates parental action demonstration in the dyadic interaction. *Developmental Science*, *18*(6), 1006–1013. https://doi.org/10.1111/desc.12270

Gabrielsson, A., & Juslin, P. N. (1996). Emotional Expression in Music Performance: Between the Performer's Intention and the Listener's Experience. *Psychology of Music*, *24*(1), 68–91. https://doi.org/10.1177/0305735696241007

Gabrielsson, A., & Lindström, E. (2010). The role of structure in the musical expression of emotions. *Handbook of music and emotion: Theory, research, application*s, *367400,* 367-400.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, *415*(6873), 755–755. https://doi.org/10.1038/415755a

Gieseking, W., & Leimer, K. (1972). *Piano Technique*. Dover Publications.

Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2015). Early development of turn-taking in vocal interaction between mothers and infants. *Frontiers in psychology*, *6*(1167), 236-245. https://doi.org/10.3389/fpsyg.2015.01167

Gruber, T., Deschenaux, A., Frick, A., & Clément, F. (2019). Group Membership Influences More Social Identification Than Social Learning or Overimitation in Children. *Child Development*, *90*(3), 728–745. https://doi.org/10.1111/cdev.12931

Gweon, H. (2020). The role of communication in acquisition, curation, and transmission of culture. *Behavioral and Brain Sciences*, *43*. https://doi.org/10.1017/S0140525X19002863

Heyes, C. (2017). When does social learning become cultural learning? *Developmental Science*, *20*(2), e12350. https://doi.org/10.1111/desc.12350

Hickey, M. (2009). Can improvisation be 'taught'?: A call for free improvisation in our schools. *International journal of music education, 27*(4), 285-299. https://doi.org/10.1177/0255761409345442

Hickey, M. (2015). Learning from the experts: A study of free-improvisation pedagogues in university settings. *Journal of Research in Music Education, 62*(4), 425-445.

https://doi.org/10.1177/0022429414556319

Ho, M. K., Littman, M., MacGlashan, J., Cushman, F., & Austerweil, J. L. (2016). Showing versus doing: Teaching by demonstration. *Advances in neural information processing systems 29*. Curran Associates, Inc.

Jacob, P., & Jeannerod, M. (2005). The motor theory of social cognition: A critique. *Trends in Cognitive Sciences*, *9*(1), 21–25. https://doi.org/10.1016/j.tics.2004.11.003

Johnson, C. M. (2000). Effect of Instruction in Appropriate Rubato Usage on the Onset Timings of Musicians in Performances of Bach. *Journal of Research in Music Education*, *48*(1), 78–83. https://doi.org/10.2307/3345458

Juslin, P. N. (2003). Five Facets of Musical Expression: A Psychologist's Perspective on Music Performance. *Psychology of Music*, *31*(3), 273–302. https://doi.org/10.1177/03057356030313003

Kallio, H., Pietilä, A. M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, *72*(12), 2954–2965. https://doi.org/10.1111/jan.13031

Király, I., Csibra, G., & Gergely, G. (2013). Beyond rational imitation: Learning arbitrary means actions from communicative demonstrations. *Journal of Experimental Child Psychology*, *116*(2), 471–486. https://doi.org/10.1016/j.jecp.2012.12.003

Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences*, *38*. https://doi.org/10.1017/S0140525X14000090

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: theory and data. *Psychology of learning and motivation*, *54*, 59-101.

Koterba, E. A., & Iverson, J. M. (2009). Investigating motionese: The effect of infant-directed action on infants' attention and object exploration. *Infant Behavior and Development*, *32*(4), 437–444. https://doi.org/10.1016/j.infbeh.2009.07.003

Laukka, P. (2004). Instrumental music teachers' views on expressivity: A report from music conservatoires. *Music Education Research*, *6*(1), 45–56. https://doi.org/10.1080/1461380032000182821

Lebler, D. (2008). Popular music pedagogy: Peer learning in practice. *Music education research, 10*(2), 193-213.

Legare, C. H., & Nielsen, M. (2015). Imitation and Innovation: The Dual Engines of Cultural Learning. *Trends in Cognitive Sciences*, *19*(11), 688–699. https://doi.org/10.1016/j.tics.2015.08.005

Lehmann, A. C., & Ericsson, K. A. (1997). Research on expert performance and deliberate practice: Implications for the education of amateur musicians and music students. *Psychomusicology: A Journal of Research in Music Cognition*, *16*, 40–58. https://doi.org/10.1037/h0094068

Levinson, S. C. (2016). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, *20*(1), 6–14. https://doi.org/10.1016/j.tics.2015.10.010

Li, S., & Timmers, R. (2020). Exploring pianists' embodied concepts of piano timbre: An interview study. *Journal of New Music Research*, *49*(5), 477–492. https://doi.org/10.1080/09298215.2020.1826532

Li, S., & Timmers, R. (2021). Teaching and learning of piano timbre through teacher–student interactions in lessons. *Frontiers in psychology*, *12*, 576056. https://doi.org/10.3389/fpsyg.2021.576056

Lindström, E., Juslin, P. N., Bresin, R., & Williamon, A. (2003). "Expressivity comes from within your soul": A questionnaire study of music students' perspectives on expressivity. *Research Studies in Music Education*, *20*(1), 23–47. https://doi.org/10.1177/1321103X030200010201

Lobel, E. E. (2021). A Model for Integrating Social and Creative Dance Forms in a Dance Education Course. *Journal of Dance Education*, *21*(1), 44–47. https://doi.org/10.1080/15290824.2020.1732984

Lombao, D., Guardiola, M., & Mosquera, M. (2017). Teaching to make stone tools: New experimental evidence supporting a technological hypothesis for the origins of language. *Scientific Reports*, *7*(1), 14394. https://doi.org/10.1038/s41598-017-14322-y

Lourenço, S. (2010). European Piano Schools: Russian, German and French classical piano interpretation and technique. *Journal of Science and Technology of the Arts / Revista de Ciência e Tecnologia Das Artes*, *2*(2-2010), 6–14. https://doi.org/10.7559/citarj.v2i1.7

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, *104*(50), 19751–19756. https://doi.org/10.1073/pnas.0704452104

Mainwaring, L. M., & Krasnow, D. H. (2010). Teaching the Dance Class: Strategies to Enhance Skill Acquisition, Mastery and Positive Self-Image. *Journal of Dance Education*, *10*(1), 14–21. https://doi.org/10.1080/15290824.2010.10387153

Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental Brain Research*, *211*(3), 547–556. https://doi.org/10.1007/s00221-011-2649-4

Marchand, T. H. J. (2010). Embodied cognition and communication: Studies with British fine woodworkers. *Journal of the Royal Anthropological Institute*, *16*(1), S100–S120. https://doi.org/10.1111/j.1467-9655.2010.01612.x

McEllin, L., Knoblich, G., & Sebanz, N. (2017). Distinct kinematic markers of demonstration and joint action coordination? Evidence from virtual xylophone playing. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(6), 885. https://doi.org/10.1037/xhp0000505

McEllin, L., Knoblich, G., & Sebanz, N. (2018). Imitation from a joint action perspective. *Mind & Language*, *33*(4), 342–354. https://doi.org/10.1111/mila.12188

McEllin, L., Sebanz, N., & Knoblich, G. (2018). Identifying others' informative intentions from movement kinematics. *Cognition*, *180*, 246–258. https://doi.org/10.1016/j.cognition.2018.08.001

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., … Glowacki, L. (2019). Universality and diversity in human song. *Science, 366*(6468). https://doi.org/10.1126/science.aax0868

Meissner, H. (2021). Theoretical framework for facilitating young musicians' learning of expressive performance. *Frontiers in psychology*, *11*, 584171. https://doi.org/10.3389/fpsyg.2020.584171

Mermelshtine, R. (2017). Parent-child learning interactions: A review of the literature on scaffolding. *British Journal of Educational Psychology*, *87*(2), 241–254. https://doi.org/10.1111/bjep.12147

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution? *Proceedings of the Royal Society B: Biological Sciences*, *285*(1880), 20180712.

https://doi.org/10.1098/rspb.2018.0712

Morgan, T. J. H., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., … Laland, K. N. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications*, *6*(1), 6029. https://doi.org/10.1038/ncomms7029

Moro, L., Mortimer, E. F., & Tiberghien, A. (2020). The use of social semiotic multimodality and joint action theory to describe teaching practices: Two cases studies with experienced teachers. *Classroom Discourse*, *11*(3), 229–251. https://doi.org/10.1080/19463014.2019.1570528

Nakamura, T. (1987). The communication of dynamics between musicians and listeners through musical performance. *Perception & psychophysics*, *41*, 525-533. https://doi.org/10.3758/BF03210487

Okazaki, S., Muraoka, Y., & Osu, R. (2019). Teacher-learner interaction quantifies scaffolding behaviour in imitation learning. *Scientific Reports*, *9*(1), 1–13. https://doi.org/10.1038/s41598-019-44049-x

Okumura, Y., Kanakogi, Y., Kobayashi, T., & Itakura, S. (2020). Ostension affects infant learning more than attention. *Cognition*, *195*, 104082. https://doi.org/10.1016/j.cognition.2019.104082

Over, H., & Carpenter, M. (2013). The Social Side of Imitation. *Child Development Perspectives*, *7*(1), 6–11. https://doi.org/10.1111/cdep.12006

Pan, Y., Novembre, G., Song, B., Li, X., & Hu, Y. (2018). Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song. *NeuroImage*, *183*, 280–290. https://doi.org/10.1016/j.neuroimage.2018.08.005

Pezzulo, G., & Dindo, H. (2011). What should I do next? Using shared representations to solve interaction problems. *Experimental Brain Research*, *211*(3), 613–630.

https://doi.org/10.1007/s00221-011-2712-1

Pezzulo, G., Donnarumma, F., & Dindo, H. (2013). Human Sensorimotor Communication: A Theory of Signaling in Online Social Interactions. *PloS one*, *8*(11), e79876. https://doi.org/10.1371/journal.pone.0079876

Pezzulo, G., Donnarumma, F., Dindo, H., D'Ausilio, A., Konvalinka, I., & Castelfranchi, C. (2019). The body talks: Sensorimotor communication and its brain and kinematic signatures. *Physics of Life Reviews*, *28*, 1–21. https://doi.org/10.1016/j.plrev.2018.06.014

Phillips-Silver, J., & Keller, P. E. (2012). Searching for roots of entrainment and joint action in early musical interactions. *Frontiers in human neuroscience*, *6*, 26. https://doi.org/10.3389/fnhum.2012.00026

Repp, B. H. (1992). A constraint on the expressive timing of a melodic gesture: Evidence from performance and aesthetic judgment. *Music Perception*, *10*(2), 221-241. https://doi.org/10.2307/40285608

Repp, B. H. (1997). Expressive timing in a Debussy Prelude: A comparison of student and expert pianists. *Musicae Scientiae*, *1*(2), 257-268. https://doi.org/10.1177/102986499700100205

Rice, T. (2003). The ethnomusicology of music learning and teaching. In *College Music Symposium* (Vol. 43, pp. 65-85). College Music Society.

Rivière, J. P., Alaoui, S. F., Caramiaux, B., & Mackay, W. E. (2019). Capturing Movement Decomposition to Support Learning and Teaching in Contemporary Dance. *Proceedings of the ACM on Human-Computer Interaction*, *3*(CSCW), 86:1-86:22. https://doi.org/10.1145/3359188

Saffran, J. R., & Kirkham, N. Z. (2018). Infant Statistical Learning. *Annual Review of Psychology*, *69*(1), 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Saint-Georges, C., Chetouani, M., Cassel, R., Apicella, F., Mahdhaoui, A., Muratori, F., … Cohen, D. (2013). Motherese in Interaction: At the Cross-Road of Emotion and Cognition? (A Systematic Review). *PloS one*, *8*(10), e78103. https://doi.org/10.1371/journal.pone.0078103

Sartori, L., Becchio, C., & Castiello, U. (2011). Cues to intention: The role of movement information. *Cognition*, *119*(2), 242–252. https://doi.org/10.1016/j.cognition.2011.01.014

Savage, P. E. (2019). Cultural evolution of music. *Palgrave Communications*, *5*(1), 1–12. https://doi.org/10.1057/s41599-019-0221-1

Schaik, J. E. van, Meyer, M., Ham, C. R. van, & Hunnius, S. (2019). Motion tracking of parents' infant- versus adult-directed actions reveals general and action-specific modulations. *Developmental Science*, *23*(1), e12869. https://doi.org/10.1111/desc.12869

Schiavio, A., Biasutti, M., van der Schyff, D., & Parncutt, R. (2020). A matter of presence: A qualitative study on teaching individual and collective music classes. *Musicae Scientiae*, *24*(3), 356–376. https://doi.org/10.1177/1029864918808833

Schiavio, A., Ryan, K., Moran, N., van der Schyff, D., & Gallagher, S. (2021). By myself but not alone: Agency, creativity, and extended musical historicity. *Journal of the Royal Musical Association*.

Schiavio, A., van der Schyff, D., Philippe, R. A., & Biasutti, M. (2022). Music teachers' self-reported views of creativity in the context of their work. *Arts and Humanities in Higher Education*, 14740222221125640. https://doi.org/10.1177/14740222221125640

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*(2), 70–76.

https://doi.org/10.1016/j.tics.2005.12.009

Sebanz, N., & Knoblich, G. (2021). Progress in Joint-Action Research. *Current Directions in Psychological Science*, *30*(2), 138–143. https://doi.org/10.1177/0963721420984425

Sebanz, N., & Shiffrar, M. (2009). Detecting deception in a bluffing body: The role of expertise. *Psychonomic Bulletin and Review*, *16*(1), 170–175. Scopus. https://doi.org/10.3758/PBR.16.1.170

Shipton, C., & Nielsen, M. (2015). Before Cumulative Culture. *Human Nature*, *26*(3), 331–345. https://doi.org/10.1007/s12110-015-9233-8

Simones, L. (2019). A framework for studying teachers' hand gestures in instrumental and vocal music contexts. *Musicae Scientiae*, *23*(2), 231–249. https://doi.org/10.1177/1029864917743089

Simones, L., Schroeder, F., & Rodger, M. (2015). Categorizations of physical gesture in piano teaching: A preliminary enquiry. *Psychology of Music*, *43*(1), 103–121. https://doi.org/10.1177/0305735613498918

Sloboda, J. A. (2000). Individual differences in music performance. *Trends in Cognitive Sciences*, *4*(10), 397–403. https://doi.org/10.1016/S1364-6613(00)01531-X

Speer, D. R. (1994). An Analysis of Sequential Patterns of Instruction in Piano Lessons. *Journal of Research in Music Education*, *42*(1), 14–26. https://doi.org/10.2307/3345333

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition, 2nd ed*. Malden: Blackwell Publishing.

Stout D. (2011). Stone toolmaking and the evolution of human culture and cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*(1567), 1050–1059. https://doi.org/10.1098/rstb.2010.0369

Strauss, S., & Ziv, M. (2012). Teaching Is a Natural Cognitive Ability for Humans. *Mind,* *Brain, and Education*, *6*(4), 186–196. https://doi.org/10.1111/j.1751-228X.2012.01156.x

Tennie, C., Call, J., & Tomasello, M. (2009). Ratcheting up the ratchet: On the evolution of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1528), 2405–2415. https://doi.org/10.1098/rstb.2009.0052

Thornton, A., & Raihani, N. J. (2008). The evolution of teaching. *Animal Behaviour*, *75*(6), 1823–1836. https://doi.org/10.1016/j.anbehav.2007.12.014

Tomasello, M. (2016). The ontogeny of cultural learning. *Current Opinion in Psychology*, *8*, 1–4. https://doi.org/10.1016/j.copsyc.2015.09.008

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Sciences*, *16*(3), 495–511. https://doi.org/10.1017/S0140525X0003123X

Tominaga, A., Knoblich, G., & Sebanz, N. (2022). Expert pianists make specific exaggerations for teaching. *Scientific Reports*, *12*(1), 21296. https://doi.org/10.1038/s41598-022-25711-3

Trehub, S. E. (2003). The developmental origins of musicality. *Nature neuroscience*, *6*(7), 669-673. https://doi.org/10.1038/nn1084

Ugur, E., Nagai, Y., Celikkanat, H., & Oztop, E. (2015). Parental scaffolding as a bootstrapping mechanism for learning grasp affordances and imitation skills. *Robotica*, *33*(5), 1163–1180. https://doi.org/10.1017/S0263574714002148

Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. *Speech Communication*, *49*(1), 2–7. https://doi.org/10.1016/j.specom.2006.10.003

Van der Schyff, D., Schiavio, A., & Elliott, D. J. (2016). Critical ontology for an enactive music pedagogy. *Action, Criticism, and Theory for Music Education, 15*(5). http://dx.doi.org/10.22176/act15.5.81

Veissière, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2020). Thinking through other minds: A variational approach to cognition and culture. *Behavioral and Brain Sciences*, *43*. https://doi.org/10.1017/S0140525X19001213

Vesper, C., Morisseau, T., Knoblich, G., & Sperber, D. (2021). When is ostensive communication used for joint action? *Cognitive Semiotics*, *14*(2), 101–129. https://doi.org/10.1515/cogsem-2021-2040

Vesper, C., & Sevdalis, V. (2020). Informing, coordinating, and performing: A perspective on functions of sensorimotor communication. *Frontiers in Human Neuroscience*, *14*, 168. https://doi.org/10.3389/fnhum.2020.00168

Volioti, G., & Williamon, A. (2017). Recordings as learning and practising resources for performance: Exploring attitudes and behaviours of music students and professionals. *Musicae Scientiae*, *21*(4), 499–523. https://doi.org/10.1177/1029864916674048

Watanabe, K. (2013). Teaching as a Dynamic Phenomenon with Interpersonal Interactions. *Mind, Brain, and Education*, *7*(2), 91–100. https://doi.org/10.1111/mbe.12011

Whiten, A. (2017). Social Learning and Culture in Child and Chimpanzee. *Annual Review of Psychology*, *68*(1), 129–154. https://doi.org/10.1146/annurev-psych-010416-044108

Wisniewski, W. W. (2016). *DEFINING NATIONAL PIANO SCHOOLS Perceptions and challenges* (Doctoral dissertation, The University of Sydney). http://hdl.handle.net/2123/14345

Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 593–602. https://doi.org/10.1098/rstb.2002.1238

Woody, R. H. (2000). Learning Expressivity in Music Performance: An Exploratory Study. *Research Studies in Music Education*, *14*(1), 14–23.

https://doi.org/10.1177/1321103X0001400102

Woody, R. H. (2003). Explaining Expressive Performance: Component Cognitive Skills in an Aural Modeling Task. *Journal of Research in Music Education*, *51*(1), 51–63. https://doi.org/10.2307/3345648

Woody, R. H. (2006). The Effect of Various Instructional Conditions on Expressive Music Performance. *Journal of Research in Music Education*, *54*(1), 21–36. https://doi.org/10.1177/002242940605400103

Wood, D., & Middleton, D. (1975). A Study of Assisted Problem-Solving. *British Journal of Psychology*, *66*(2), 181–191. https://doi.org/10.1111/j.2044-8295.1975.tb01454.x

Wu, H. G., Miyamoto, Y. R., Castro, L. N. G., Ölveczky, B. P., & Smith, M. A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience*, *17*(2), 312–321. https://doi.org/10.1038/nn.3616

Yoon, J. M. D., Johnson, M. H., & Csibra, G. (2008). Communication-induced memory biases in preverbal infants. *Proceedings of the National Academy of Sciences*, *105*(36), 13690–13695. https://doi.org/10.1073/pnas.0804388105

Zamm, A., Wang, Y., & Palmer, C. (2018). Musicians' Natural Frequencies of Performance Display Optimal Temporal Stability. *Journal of Biological Rhythms*, *33*(4), 432–440. https://doi.org/10.1177/0748730418783651

Zukow-Goldring, P., & Arbib, M. A. (2007). Affordances, effectivities, and assisted imitation: Caregivers and the directing of attention. *Neurocomputing*, *70*(13), 2181–2193. https://doi.org/10.1016/j.neucom.2006.02.029