How do expert pianists adapt their pedagogical demonstration to novices' skill levels?

Atsuko Tominaga[1], Günther Knoblich[1], & Natalie Sebanz[1]

[1] Department of Cognitive Science, Central European University

Author Note

Correspondence concerning this article should be addressed to Atsuko Tominaga, Quellenstraße 51, 1100 Vienna, Austria. E-mail: tominaga_atsuko@phd.ceu.edu

Abstract

Later

Keywords: keywords

How do expert pianists adapt their pedagogical demonstration to novices' skill levels?

Introduction

Teaching is effective when experts and novices dynamically interact with each other (Byrne & Rapaport, 2011). In the literature on cultural evolution, it has been considered that humans have developed knowledge and skills by imitating knowledgeable others (experts) and transmitting what they have acquired to the next generation (Mesoudi, 2016). However, information flows between experts and novices are mostly not unidirectional (i.e., from experts to novices) but rather bidirectional and interactive (Charbonneau & Strachan, 2022). Experts adapt their behaviour to novices and scaffold the acquisition process of complex skills, and novices are actively participating in the learning process (Flynn, 2022; Okazaki, Muraoka, & Osu, 2019).

Exaggeration of movement is often used to signal a communicative intent to others during real-time interactions (Pezzulo et al., 2019). When experts have an intention to teach, they modulate their behaviour. For instance, when caregivers were teaching how to use a novel toy to their infants, they tended to exaggerate the movement paths (Brand, Baldwin, & Ashburn, 2002). Such exaggeration was observed not only towards infant learners but also when adults were interacting with adult learners (McEllin, Knoblich, & Sebanz, 2017; Uther, Knoll, & Burnham, 2007).

Furthermore, recent evidence has shown that such pedagogical modulations were sensitive to the skill levels of learners. Fukuyama et al. (2015) revealed mothers who were demonstrating how to use nesting cups to their infants dynamically modulated their demonstration according to the infants' object manipulation. For example, the mothers exaggerated their movements when their infants failed to produce expected actions whereas the mothers did not produce such exaggeration when their infants did not have enough motor skills to produce such actions. The results indicated that adults monitored infants' actions and abilities, and adapted their demonstration accordingly.

In our previous research (Tominaga et al., under review), we demonstrated that expert pianists modulated their performance when they were asked to teach musical expressive techniques such as articulation and dynamics. For example, expert pianists exaggerated dynamics contrast (i.e., making larger contrast between forte and piano) when they had an intention to teach dynamics. Such exaggeration is considered to be most effective for those who fail to produce the techniques in their performance. If teaching is an adaptive and scaffolding process, it can be expected that experts should not use such exaggeration if novices became able to produce them.

In the current study, we investigated whether and how expert pianists adapted their performance depending on the skill levels of novice learners. Instead of recruiting actual novice pianists, we created artificial recordings to manipulate the skill levels of novices to maximise the experimental control. The recordings were made based on the performance data from our previous experiment (Tominaga et al., under review). All the recordings had the same musical structure (i.e., the same piece) but had different implementations of articulation and dynamics. In the ideal performance, both articulation and dynamics should be implemented as notated on the sheet music (*Fig 1*). We made the recordings so that both techniques were implemented in the first quarter, only articulation was implemented in the second quarter, only dynamics was implemented in the third quarter and neither of them was implemented in the last quarter of the recordings. Therefore, in this experiment, the skill levels of novices corresponded to the implementation of articulation and dynamics.

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.

Methods

Participants

We recruited 20 participants 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 the Psychological Research Ethics Board (PREBO) CEU PU in Austria.

Apparatus and stimuli

The experiment was programmed in Max/MSP (8.1.11; https://cycling74.com/products/max) on a Mac Book 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. 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.

As stimuli, one piece of music, which was used in our previous experiment, was selected (Tominaga et al., under review). The piece was taken 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 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 *Fig 1*.

We created artificial recordings of novices to manipulate their skill levels. In this experiment, we made the recordings by changing two factors Articulation (present or absent) and Dynamics (present or absent). Therefore, there were four types of recordings. We aimed to create a quarter of the recordings where both articulation and dynamics were implemented. The second quarter included the recordings where only articulation was implemented whereas dynamics was missing. The third quarter included the recordings where only dynamics was implemented whereas articulation was missing. The last quarter included the recordings where neither articulation nor dynamics was implemented. We generated 4 instances for each type, therefore there were 16 stimuli (i.e., recordings) in total. How we generated the recordings is written in *Supplementary Material*.

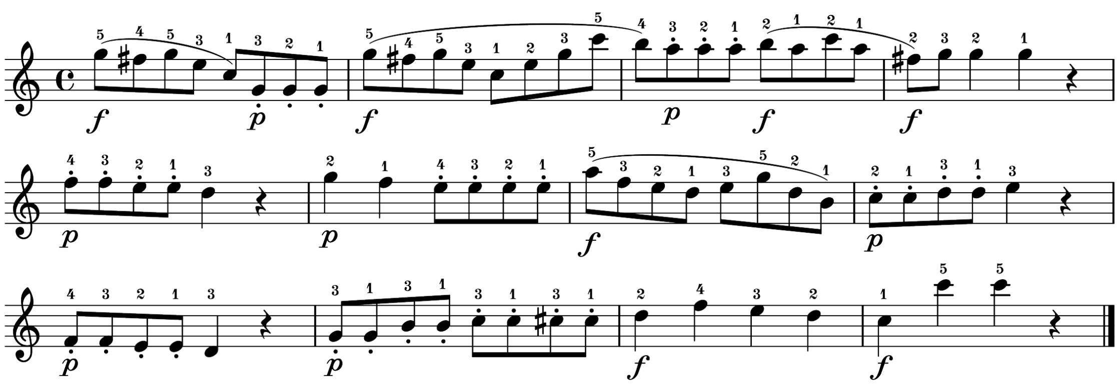

Figure 1. Sheet music. For articulation notation, the curved line (slur) indicates legato and the dots indicate staccato. For dynamics notation, the symbol 'f' denotes forte and the symbol 'p' denotes piano. For data analysis, only the 8th notes with expressive notations were included.

Procedure

Prior to the experiment, participants were required to memorise the piece so that they had enough time to practise and perform it without pitch errors while implementing notated expressions in the experiment.

First, we recorded participants' baseline performance by asking them to perform the piece without listening 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 1*) was displayed in front of the participants. Also, they were told to perform the piece expressively with their interpretation and do their best as a performer. This instruction was given to make sure that they paid attention to the expressive aspects of the performance. Each participant performed the piece twice.

After we recorded the participants' baseline performance, participants were informed 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 the student (i.e., recording). In total, there were 16 trials and participants played the piece for each student (i.e., recording) 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 in the same situation as we recorded their baseline performance. At the end of the experiment, participants filled in a questionnaire asking about their demographic information and experience in piano performance/teaching.

Data analysis

The dependent variables were computed from MIDI data for data analysis. Interonset intervals (IOIs) are the intervals between onsets of adjacent notes and provide a measure of tempo. 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 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 note. 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). 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.

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 either, 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 corrected.

After removing pitch errors, we removed outliers for IOIs, KOT, KV and KV Difference, 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.

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

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).

IOIs

Neither main effect of Articulation ($F(1, 19) = 2.79$, $p = 0.111$, $\eta^2_G = 0.010$) or Dynamics ($F(1, 19) = 3.27$, $p = 0.086$, $\eta^2_G = 0.003$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 1.48$, $p = 0.24$, $\eta^2_G = 0.002$). Therefore, participants did not change the tempo depending on the types of recordings (*Fig 2*).

KOT

Legato. There was a significant main effect of Articulation ($F(1, 19) = 5.59$, $p = 0.029$, $\eta^2_G = 0.002$). However, there was no significant main effect of Dynamics ($F(1, 19) = 0.06$, $p = 0.81$, $\eta^2_G = 0.000$) or interaction between Articulation and Dynamics ($F(1, 19) = 1.89$, $p = 0.19$, $\eta^2_G = 0.002$). Participants produced longer legato when listening to the recordings where articulation was not implemented (*Fig 3*, left).

Staccato. There was a significant main effect of Articulation ($F(1, 19) = 4.88$, $p = 0.040$, $\eta^2_G = 0.009$). However, there was no significant main effect of Dynamics ($F(1, 19) = 0.68$, $p = 0.42$, $\eta^2_G = 0.000$) or interaction between Articulation and Dynamics ($F(1, 19) = 3.13$, $p = 0.093$, $\eta^2_G = 0.001$). Participants produced shorter staccato when listening to the recordings where articulation was not implemented (*Fig 3*, right).

KV

Forte. Neither main effect of Articulation ($F(1, 19) = 1.66$, $p = 0.213$, $\eta^2_G = 0.00$) or Dynamics ($F(1, 19) = 0.33$, $p = 0.57$, $\eta^2_G = 0.000$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 0.00$, $p = 0.96$, $\eta^2_G = 0.000$). Participants did not change performances in terms of notated forte depending on the types of recordings (*Fig 4*, left).

Piano. There was a significant main effect of Articulation ($F(1, 19) = 7.18$, $p = 0.01$, $\eta^2_G = 0.060$). However, there was no significant main effect of Dynamics ($F(1, 19) = 1.80$, $p = 0.20$, $\eta^2_G = 0.003$) or interaction between Articulation and Dynamics ($F(1, 19) = 0.04$, $p = 0.85$, $\eta^2_G = 0.000$). Participants produced softer piano when listening to the recordings where articulation was implemented (*Fig 4*, right).

KV Difference

Forte to Piano. Neither main effect of Articulation ($F(1, 19) = 3.40$, $p = 0.081$, $\eta^2_G = 0.012$) or Dynamics ($F(1, 19) = 1.58$, $p = 0.22$, $\eta^2_G = 0.002$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 0.97$, $p = 0.336$, $\eta^2_G = 0.001$).

Piano to Forte. Neither main effect of Articulation ($F(1, 19) = 2.25$, $p = 0.150$, $\eta^2_G = 0.005$) or Dynamics ($F(1, 19) = 3.76$, $p = 0.07$, $\eta^2_G = 0.004$) nor the interaction between Articulation and Dynamics was significant ($F(1, 19) = 1.08$, $p = 0.31$, $\eta^2_G = 0.001$).

These results indicated that participants did not change performances in terms of dynamics contrast depending on the types of recordings (*Fig 5*).

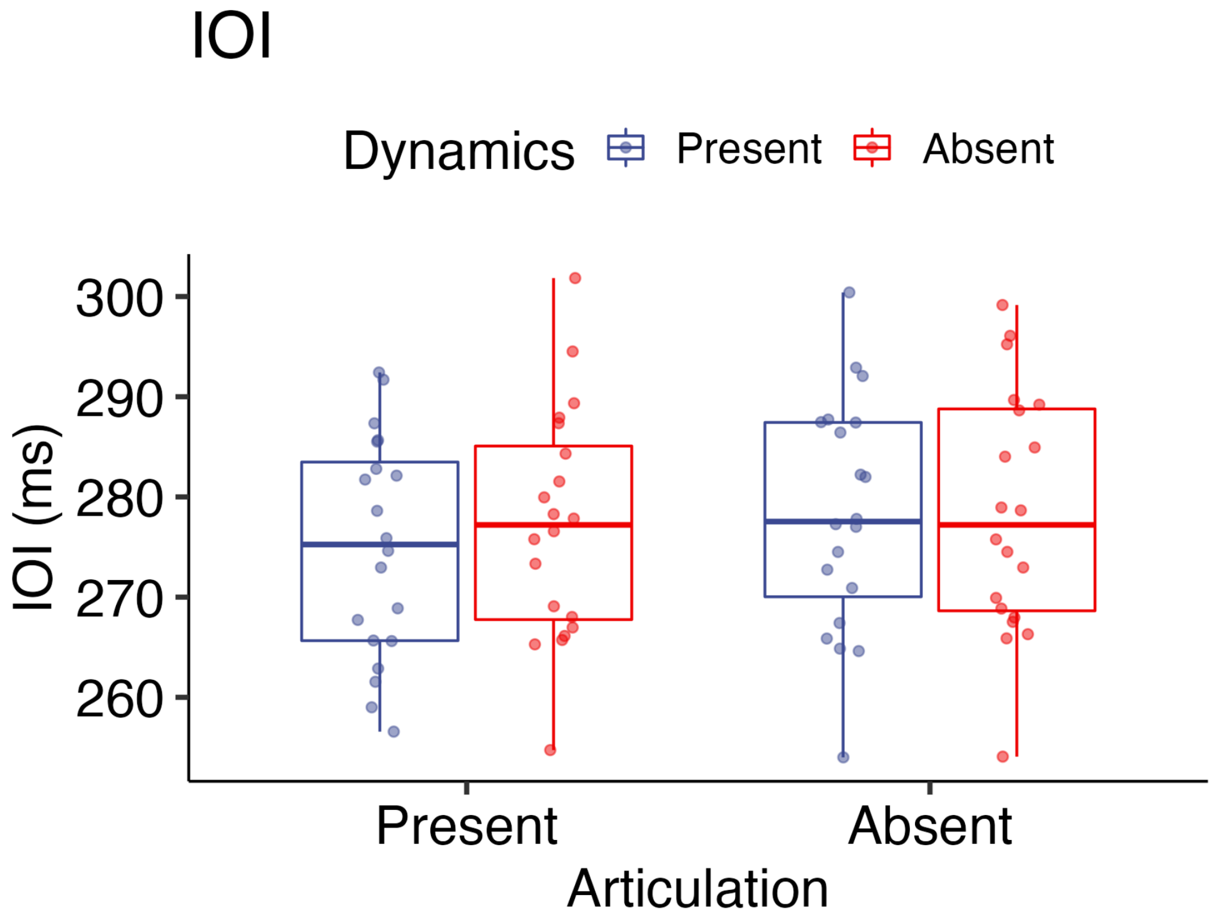

Figure 2. IOIs (ms). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

Comparison with baseline performance

In order to investigate how their performances during the experiment differed from their baseline performances (i.e., performances that participants produced before listening to any of the recordings), we performed one-way repeated-measures ANOVAs. For KOT, KV and KV Difference, we performed ANOVAs separately for each subcomponent. To recall the instruction for the baseline performance, we asked participants to perform the piece expressively with their interpretation and do their best as a performer. By comparing with the baseline performance, we can investigate how pedagogical intentions influence

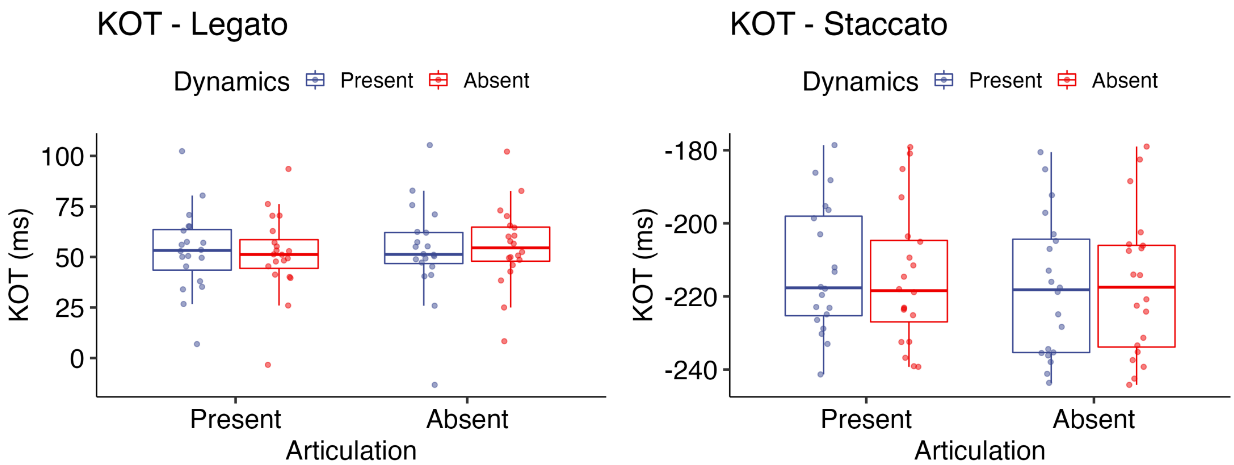

Figure 3. KOT(ms) for each subcomponent; legato (left) and staccato (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

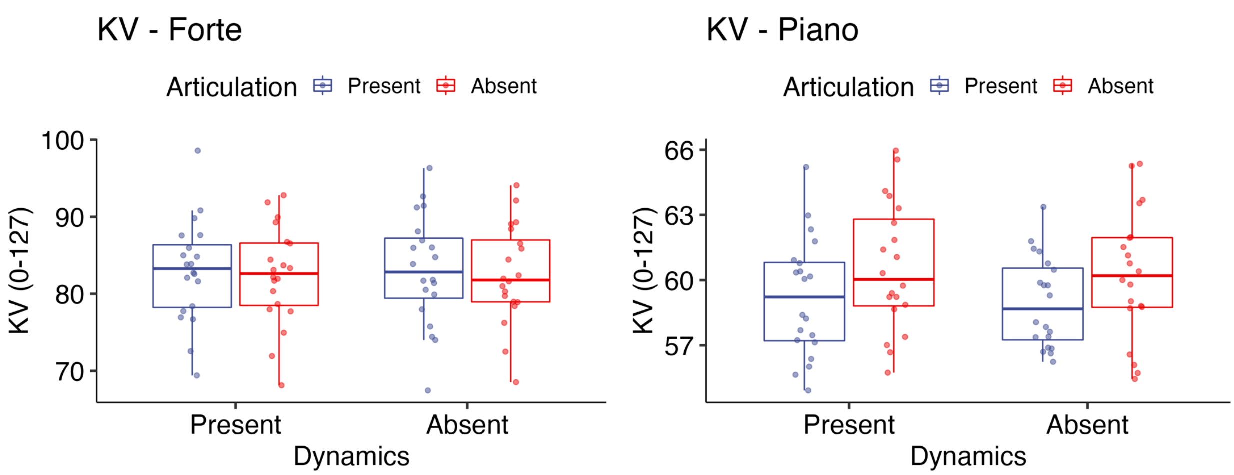

Figure 4. KV (0-127) for each subcomponent; forte (left) and piano (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

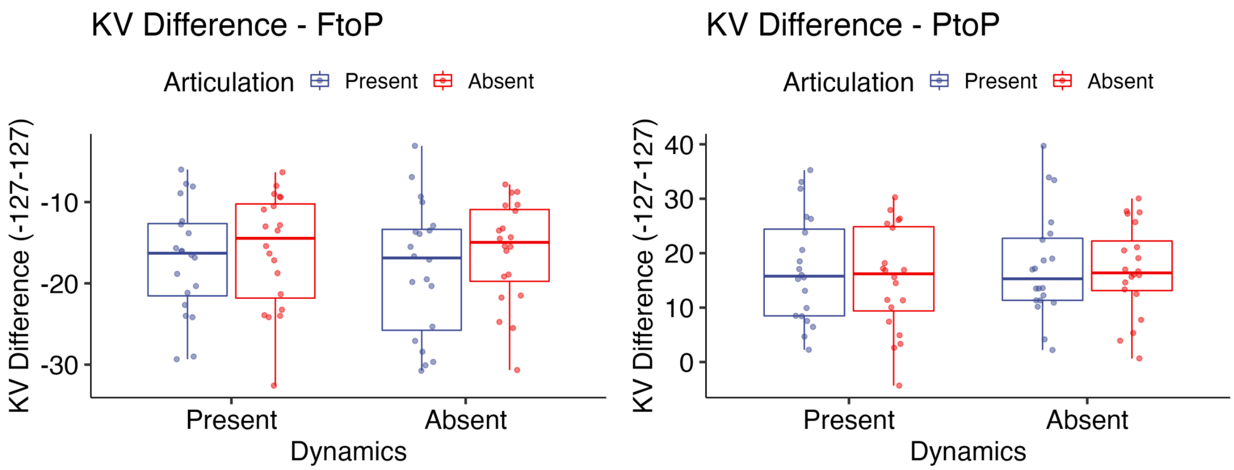

Figure 5. KV Difference (-127-127) for each subcomponent; forte to piano (left) and piano to forte (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

participants' performances.

IOIs

We compared the baseline performance with the performances where either both articulation and dynamics were implemented or neither of them was implemented to examine how participants performed differently for these extreme examples of novices' skill levels. We categorised performances into three groups (baseline, both, none) and treated them as a factor Category.

There was no significant main effect of Category ($F(1.11, 21.18) = 0.58$, $p = 0.47$, $\eta^2_G = 0.012$; Greenhouse-Geisser corrected), suggesting that participants kept the same tempo regardless of whether they had an intention to teach (*Fig 6*).

KOT

We compared the baseline performance with the performances where dynamics was implemented to examine how 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 ($F(1.45, 27.53) = 11.82$, $p < 0.001$, $\eta^2_G = 0.060$; 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 an intention to teach (*Fig 7*, left).

Staccato. There was no significant main effect of Category ($F(1.12, 21.28) = 0.75$, $p = 0.41$, $\eta^2_G = 0.011$; Greenhouse-Geisser corrected), suggesting that participants did not play staccato differently depending on whether they had an intention to teach (*Fig 7*, right).

KV

We compared the baseline performance with the performances where articulation was implemented to examine how 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 ($F(1.59, 30.18) = 1.43$, $p = 0.25$, $\eta^2_G = 0.004$; Greenhouse-Geisser corrected), suggesting that participants did not play forte differently depending on whether they had an intention to teach (*Fig 8*, left).

Piano. There was a significant main effect of Category ($F(1.79, 34.09) = 5.68$, $p = 0.009$, $\eta^2_G = 0.072$; 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 an intention to teach (*Fig 8*, right).

KV Difference

We compared the baseline performance with the performances where articulation was implemented to examine how 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 ($F(1.53, 29.15) = 1.73$, $p = 0.20$, $\eta^2_G = 0.014$; Greenhouse-Geisser corrected).

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

These results indicated that participants did not make dynamics contrast between forte and piano differently depending on whether they had an intention to teach (*Fig 9*).

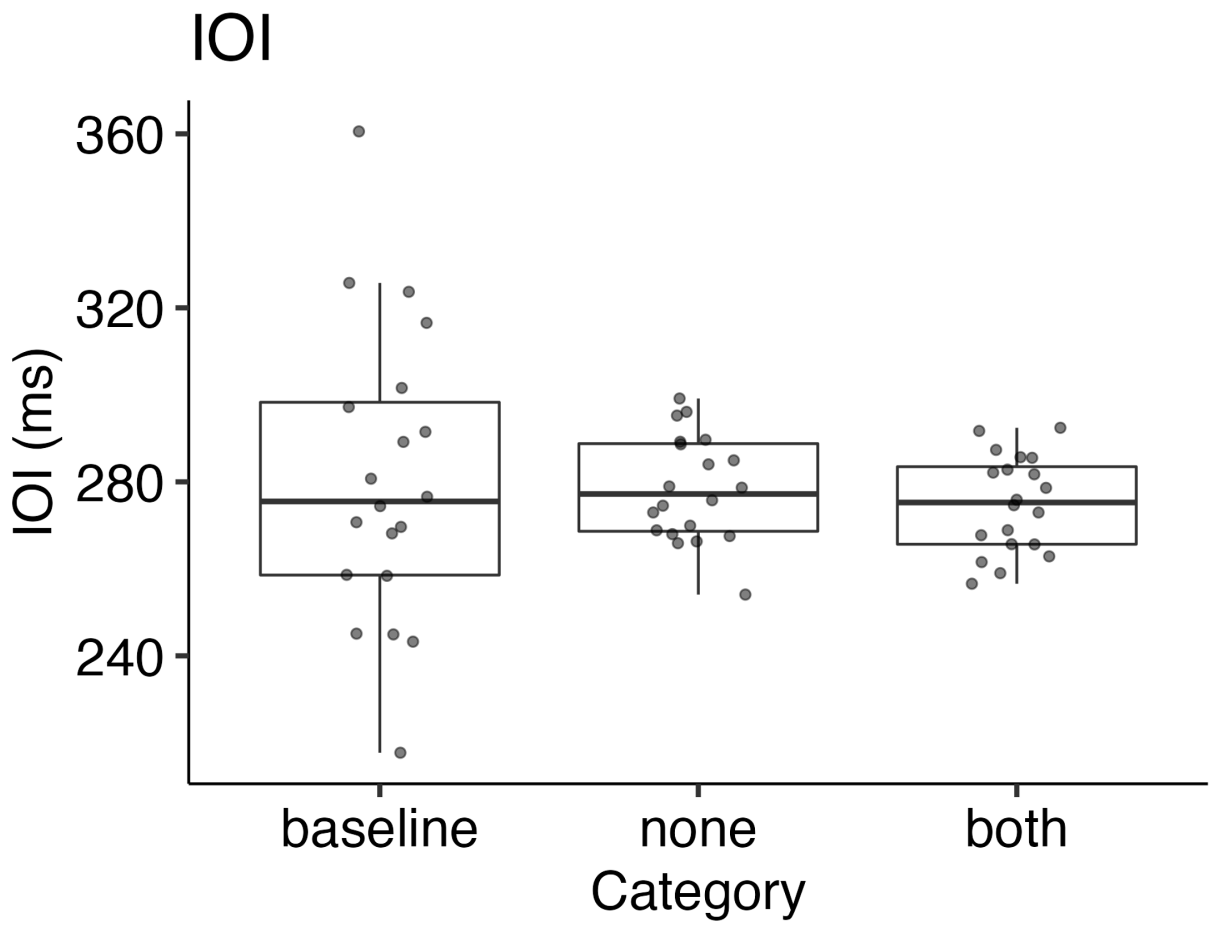

Figure 6. Comparison with the baseline performance in terms of IOIs (ms). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

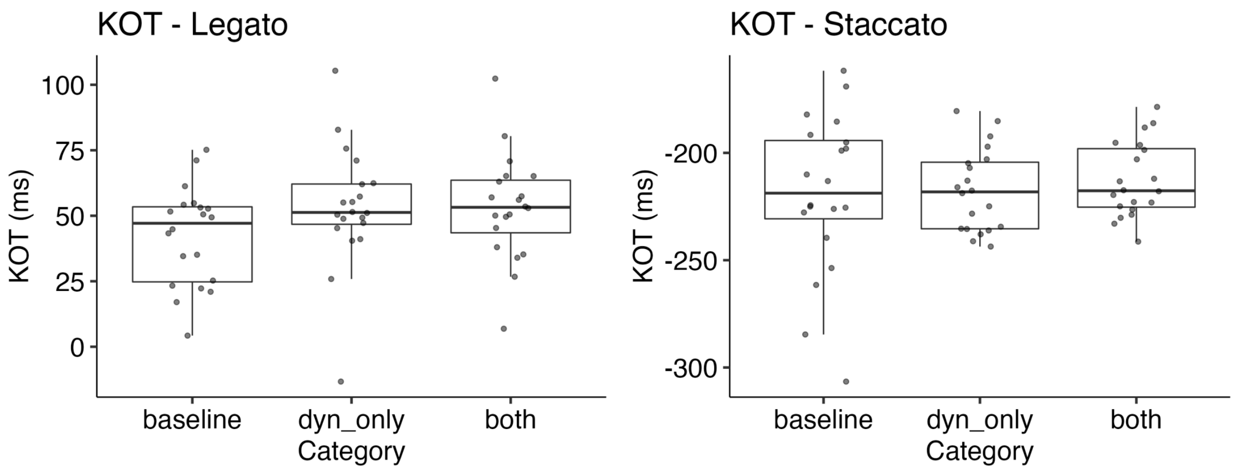

Figure 7. Comparison with the baseline performance in terms of KOT(ms) for each subcomponent; legato (left) and staccato (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

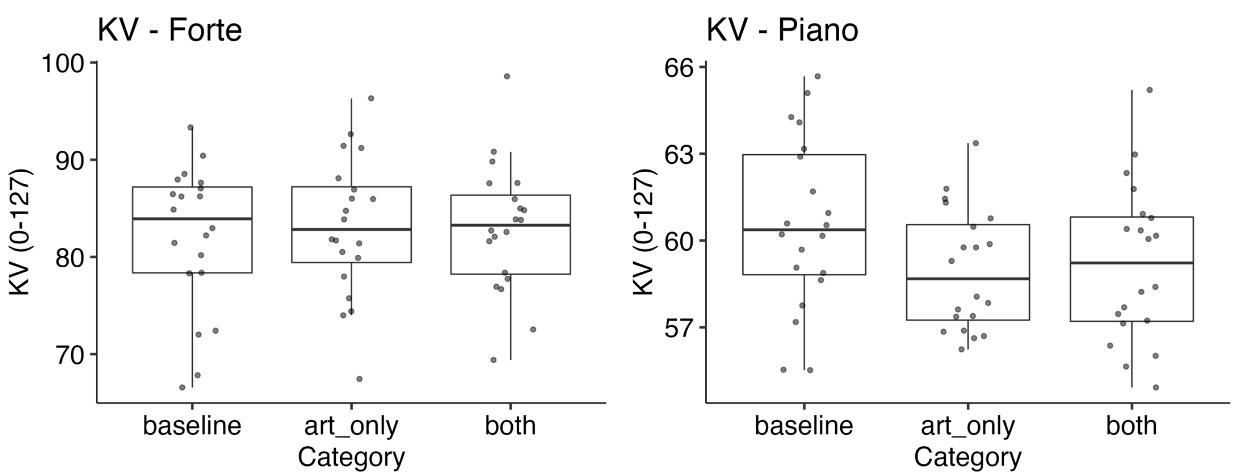

Figure 8. Comparison with the baseline performance in terms of KV (0-127) for each subcomponent; forte (left) and piano (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

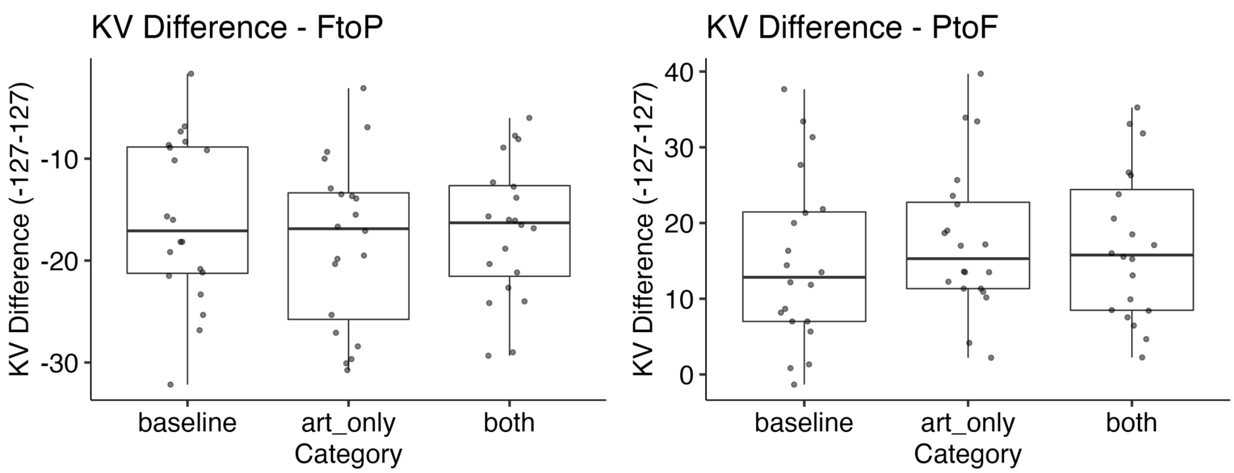

Figure 9. Comparison with the baseline performance in terms of KV Difference (-127-127) for each subcomponent; forte to piano (left) and piano to forte (right). Each box indicates the IQR with the median, and whiskers extend to a maximum of 1.5 × IQR beyond the box.

Discussion

The present study investigated whether and how expert pianists adapt their performance depending on the skill levels of novices. We created artificial recordings to manipulate the skill levels by modifying 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 replicate the performances of students with high skills. On the other hand, the recordings where none of the two techniques was implemented were supposed to replicate the performances of students with low skills. The recordings where either of the two techniques was implemented but the other was missing were equivalent to the performances of students with intermediate skills.

We found that expert pianists did not modulate their tempo (IOIs) depending on the skill levels of students. Even compared with the tempo of the baseline performance, they did not change the tempo either. These findings indicated that tempo was not employed specifically for teaching because tempo itself was relevant to neither articulation nor dynamics. Although slower demonstration is generally considered to be used for teaching (McEllin et al., 2017; Saint-Georges et al., 2013; Schaik, Meyer, Ham, & Hunnius, 2019), specially in music performance, it may not be an effective strategy to perform slowly as changing tempo might give another interpretation of music (e.g., expressive timing).

For KOT, 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.

Overall, the results indicated that expert pianists seemed to prioritise the teaching of articulation. One possibility would be that articulation was more important for the specific piece we selected or the piece itself naturally invited some implementation of articulation. In our previous experiment, we observed most participants implemented articulation (particularly legato) when they were asked to perform the piece even when neither of 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 trying to teach 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.

One of the limitations of the current study is that participants did not have an idea about what the ideal performance was. 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 create a model performance where both articulation and dynamics was implemented and ask participants (expert pianists) to listen to it before starting the current experiment 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 all the recordings did not include any pitch errors and were performed the piece with a stable tempo. This was to make sure that we manipulated articulation and dynamics features of the recordings only. However, all the recordings might not have sounded as if unskillful students were playing. In order to create more realistic unskillful performances, it might be useful to add some jitters to the tempo of the recordings we created, or to ask unskillful students to play the piece and extract the temporal feature 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). In the current experiment, expert pianists had only 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 performance of expert pianists and that of novices are related to each other.

References

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

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

Charbonneau, M., & Strachan, J. W. A. (2022). From Copying to Coordination: An Alternative Framework for Understanding Cultural Learning Mechanisms. *Journal of Cognition and Culture*, *22*(5), 451–466. https://doi.org/10.1163/15685373-12340145

Flynn, E. (2022). "Imitation Is the Sincerest Form of" ... Cultural Evolution, or Is It? *Journal of Cognition and Culture*, *22*(5), 436–450. https://doi.org/10.1163/15685373-12340144

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

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

Mesoudi, A. (2016). Cultural evolution: Integrating psychology, evolution and culture. *Current Opinion in Psychology*, *7*, 17–22. https://doi.org/10.1016/j.copsyc.2015.07.001

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

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

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

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*, *0*(0), e12869. https://doi.org/10.1111/desc.12869

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

Acknowledgement

This research was supported by the European Research Council under the European Union's Seventh Framework Program (FP7/2007–2013)/ERC Grant agreement no. 609819, SOMICS, and by ERC Grant agreement no. 616072, JAXPERTISE. We thank Candasch Acar for his help with data collection.