



Short Communication

Joint action coordination in expert-novice pairs: Can experts predict novices' suboptimal timing?

Thomas Wolf*, Natalie Sebanz, Günther Knoblich

Department of Cognitive Science, Central European University, Budapest, Hungary



ARTICLE INFO

Keywords:

Joint action
Temporal coordination
Sensorimotor synchronization
Prediction
Expertise
Music ensembles

ABSTRACT

Previous research has established that skilled joint action partners use predictive models to achieve temporal coordination, for instance, when playing a music duet. But how do joint action partners with different skill levels achieve coordination? Can experts predict the suboptimal timing of novices? What kind of information allows them to predict novices' timing? To address these questions, we asked skilled pianists to perform duets with piano novices. We varied whether, prior to performing duets, experts were familiar with novices' performances of their individual parts of the duets and whether experts had access to the musical scores including the novices' part of the duet. Familiarity with the score led to better coordination when the score implied a difficult passage. Familiarity with novices' performances led to better joint action coordination for the remaining parts of the duet. Together, the results indicate that experts are surprisingly flexible in predicting novices' suboptimal timing.

1. Introduction

When people perform joint actions together, they need to coordinate their actions in time (Butterfill, 2017; Keller, Novembre, & Hove, 2014; Pecenka & Keller, 2011; Sebanz & Knoblich, 2009). Previous research has demonstrated that internal models enable joint action partners to predict each other's timing if both are skilled in performing the individual parts of a joint action (e.g., Kourtis, Knoblich, Woźniak, & Sebanz, 2014). However, an open question is how joint action partners who differ in their individual skills achieve joint action coordination. For instance, when two musicians play a piano duet together, they need to adhere to certain tempo requirements and to minimize interpersonal asynchronies, regardless of differences in their skills. How coordination is achieved despite large differences in skill is an important question because it pertains to many joint actions performed in the context of teaching (Csibra & Gergely, 2009). The aim of the present study was to investigate whether skilled performers can ensure successful interpersonal coordination despite novices' suboptimal timing, and what kind of information helps them to achieve this.

We started from the hypothesis that interpersonal temporal coordination can be achieved if the skilled joint action partner compensates for a novice's suboptimal performance. When adapting to novices' suboptimal timing experts have to go beyond using their own internal models to predict a joint action partner's performance in real time (Noy, Dekel, & Alon, 2011; Wolpert, Doya, & Kawato, 2003). The reason is

that, according to internal model accounts, the accuracy of predictions of a joint action partner's performance should depend on a high degree of similarity of the predicting and the predicted system (Grèzes, Frith, & Passingham, 2004; Knoblich & Flach, 2001). The dimensions of similarity that can affect the accuracy of coordination range from a shared preference for a general tempo (Loehr & Palmer, 2011) to fine-grained similarities in micro-timing that characterize particular individual performances (Keller, Knoblich, & Repp, 2007). Although there is some evidence that, in the context of action observation, predictions can be adjusted to reflect general differences in action capabilities (Ramenzoni, Riley, Davis, Shockley, & Armstrong, 2008; Welsh, Wong, & Chandrasekharan, 2013), it is an open question whether temporal predictions can be adjusted to the suboptimal timing that characterizes novice performance in domains such as sports, dance, and music where such joint actions are frequent.

We studied this question in the domain of musical joint action, where continuous and accurate temporal coordination is crucial to achieve successful joint performance (Keller et al., 2014). It is well established that music experts use internal models to predict the outcomes of their own and others' actions (Haueisen & Knösche, 2001) and experts' performance parameters differ systematically from novices' performance parameters (Aoki, Furuya, & Kinoshita, 2005; Loehr & Palmer, 2007).

Observation of music ensembles rehearsals (Ginsborg, Chaffin, & Nicholson, 2006; Ginsborg & King, 2012) suggested two important sources of information for improving temporal coordination:

* Corresponding author at: Department of Cognitive Science, Central European University, Október 6 utca 7, 1051 Budapest, Hungary.
E-mail address: wolf_thomas@phd.ceu.edu (T. Wolf).

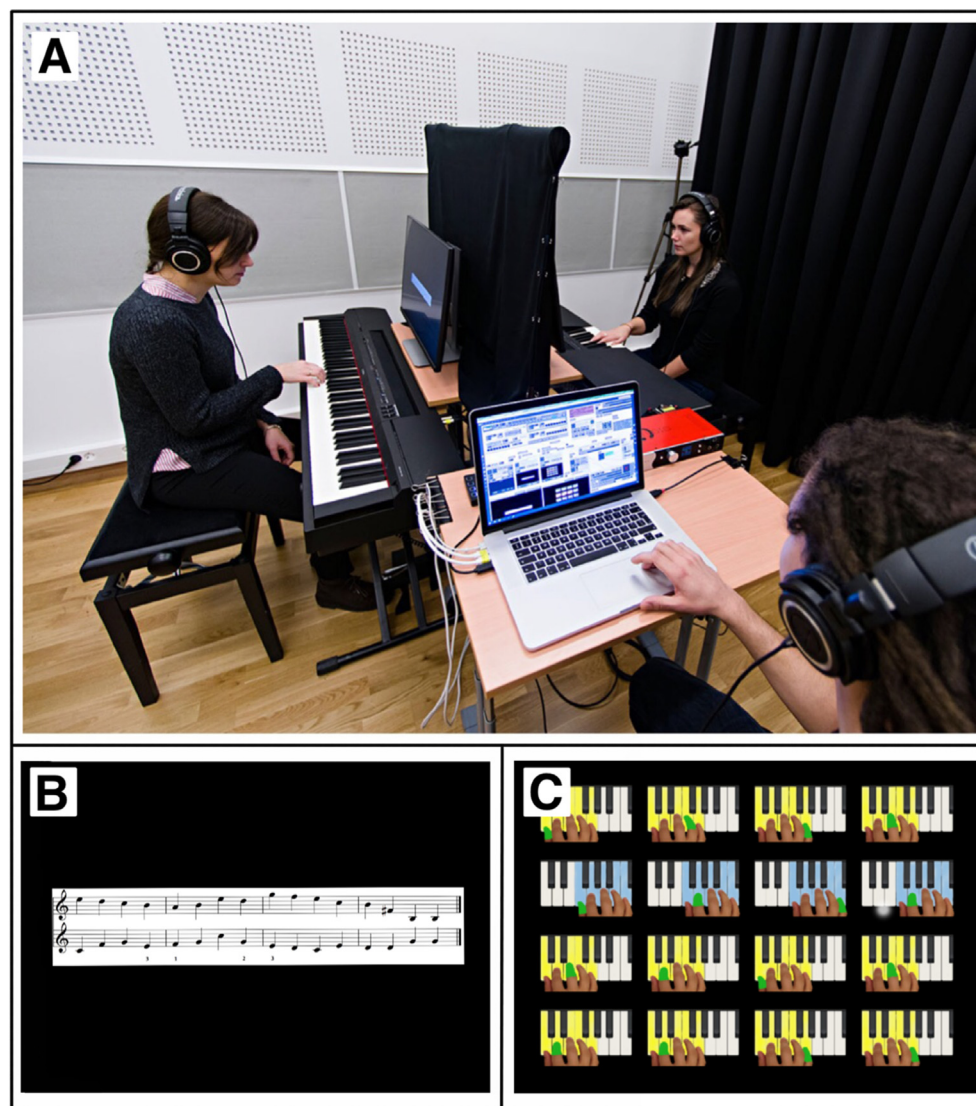


Fig. 1. (A) Experimental setup. (B) Experts performed their own part based on a standard musical notation (upper row). For half of the duets experts received the full musical score including the novice's part (lower row). (C) Novices performed their part based on a simple custom notation that was read line by line from top to bottom. Each picture corresponded to one beat and showed one of two hand positions (color-coded with yellow and blue) and the finger to be used to press one of the piano keys (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

knowledge of the structure of a partner's part, and knowledge of a partner's interpretation of her part, including expressive timing (Repp, 1990). The former can be acquired in the absence of a partner from musical scores. In contrast, a partner's idiosyncratic interpretation will need to be experienced first-hand.

A study by Ragert, Schroeder, and Keller (2013) provided support for the importance of experiencing a partner's performance. They asked highly trained pairs of piano experts to repeatedly perform duets. Temporal coordination between the pianists improved across consecutive repetitions of the same duet, supporting the claim that increasing familiarity with a partner's playing style improved interpersonal coordination. This finding indicates that pianists were able to adjust their predictions to capture parameters of an expert partner's timing. However, it is an open question whether experts are also able to adjust to novices' suboptimal timing that does not express a certain musical style. This may be a precondition for teaching through joint actions where experts provide a timing scaffold for novices.

But are there any regularities in novices' timing that experts could pick up on to improve their predictions? In order to address this question, we varied whether experts were familiar with novices'

performances of their individual parts before performing a duet with them. We hypothesized that experts' familiarity with novices' playing would improve temporal coordination during ensuing duet performance. This is only expected to occur if experts can extract from a novice's performance idiosyncratic patterns that help them to improve their real-time predictions of the novice's performance during a joint performance. One factor that is likely to produce predictable timing variability in the novice is encountering particular motor difficulties such as having to switch the hand position on an instrument. Such difficulties are often visible from the musical score because these scores can include particular instructions for the positioning of fingers and hands. In order to find out whether experts can translate such symbolic information into real-time predictions during joint performance we used music notation from the tradition of Western classical music and varied whether experts knew in advance the musical score including the novices' parts of the duets. Knowing novices' scores is only expected to help experts to improve temporal coordination during joint performance if they can identify difficulties for the novices and translate these into accurate delays in their own performance to match delays in the novice's performance. Finally, to check that timing variability in

novices originate from their performance difficulties, i.e. was sub-optimal, we attempted to rule out the theoretical possibility that some of their variability originates from expressive timing based on musical intentions. This was done by checking for autocorrelations in novices' performances which are present in experts' expressive timing.

2. Methods

2.1. Participants

Twelve expert pianists (5 women, 7 men, mean age = 24 years, $SD = 3$ years) participated in the experiment. All experts had at least 10 years of private piano lessons ($M = 12$ years, $SD = 3$ years). Twelve non-pianists (8 women, 4 men, mean age = 25 years, $SD = 5$ years) were randomly paired with the 12 experts. None of the novices had received piano lessons, but all of them had completed minimally 5 years of private lessons on another instrument ($M = 9$ years, $SD = 3$ years). Experts and novices were recruited through flyers distributed in music schools and university campuses in Budapest. All participants gave their informed consent and received gift vouchers as compensation. This study was approved by the United Ethical Review Committee for Research in Psychology (EPKEB) in Hungary.

2.2. Apparatus and material

Participants played on two Yamaha digital pianos. Presentation of visual information (standard scores and adapted scores for novices), auditory feedback, metronome beats, and data recording were implemented using a custom Max MSP patch. The eight duets the participants performed were based on easy piano duets from the standard repertoire. Expert and novice parts were simplified, shortened, and modified to conform to a length of four bars of four quarter notes each. The novice parts were adapted so that they could be performed with the white keys from C4 to C5. A special notation allowed the novices to sight-read simple melodies and to reproduce them on the piano after a short training (see Fig. 1C). For novices, each melody started in a lower hand position (thumb on C4) and required two shifts of hand position, first upwards (thumb from C4 to F4) and then downwards (thumb from F4 to C4). This means that in each melody there were two shifts to be performed by the novice. These two shifts per melody will be referred to as difficult passages, while the rest of the melody, where novices do not have to shift their hands, will be referred to as easy passages.

2.3. Procedure and design

The main experiment consisted of eight blocks. In each block, the novice performed the duet eight times. The first four repetitions were performed by the novice alone and therefore characterize individual novices' performances of their part of the duet. In the second four repetitions the expert played along with the novice. Uni-directional auditory feedback ensured that the novice did not hear the expert and thus could not adapt to the expert's performance. Experts performed their parts of the duets individually after the main experiment for two repetitions.

In the first part of each block, where novices played alone, the expert either heard the novices' four individual performances or not (Familiarity with Novice Performance, yes or no) and the expert either had access to the sheet music for the novice's part or not (Familiarity with the Score, yes or no). In the second part of the block novices never heard the experts and were instructed to simply repeat playing their parts of the duet four more times. Experts played their part of the duet along with the novice, with the instructions to synchronize their keystrokes with the novices' keystrokes as accurately as possible. An occluder prevented visual contact between the two participants of a pair (see Fig. 1A). The data for each cell in this 2×2 within-participant design were derived from two different duets (blocks).

Novices were invited to a training session that took place 2–14 days in advance of the main experiment. The aim of the training session was to ensure that novices could produce error-free performances of their parts of the duets. Each novice played the novice part of each of the eight duets eight times. As in the main experiment each performance was preceded by four leading metronome beats at a tempo of 60 bpm. Novices were instructed to match this target tempo. All twelve novices managed to reach the criterion of three error-free performances in a row for all melodies.

3. Results

Before analyzing synchronization performance, we compared the variability of Inter-Keystroke-Intervals (IKIs) of individual performances of novices and experts. As expected, novices had larger individual variability in IKIs, i.e., higher SD of IKIs ($M = 44$ ms, $SD = 7$) than experts ($M = 35$ ms, $SD = 7$). A Welch t -test revealed that this difference was significant, $t(21.97) = 3.30$, $p = .003$ (Delacre, Lakens, & Leys, 2017). Furthermore, experts' variability in IKIs during test trials was significantly higher ($M = 50$ ms, $SD = 9$), than during their individual performances ($M = 35$ ms, $SD = 7$), $t(21.08) = 4.68$, $p < .001$.

As a measure of interpersonal temporal coordination, we analyzed the asynchronies between novices' and experts' keystrokes in the test phase where the expert accompanied the novice. Of 512 asynchronies per pair, 12.5% were produced after a hand shift. A small portion of data points were excluded due to technical error (0.26%) or wrong pitches, insertions, omissions, incorrect alignment (6.24%). From the raw asynchronies, we derived the mean absolute asynchronies and the variability of absolute asynchronies. Both measures were analyzed with 2×2 repeated measures ANOVAs with the factors Familiarity with Novice Performance (yes/no) and Familiarity with Score (yes/no). Asynchronies with preceding hand shifts (Difficult Passages) and asynchronies without preceding hand shifts (Easy Passages) were analyzed separately because they reflect qualitatively different movements that imply different levels of difficulty for novices.

3.1. Mean absolute asynchronies

Panels A and B in Fig. 2 show the results for mean absolute asynchronies in ms separately for Easy and Difficult Passages. For keystrokes from Easy Passages (Panel A) the mean asynchrony between expert and novice was lower when the expert had heard the novice practice her melody beforehand ($M = 44$ ms, $SD = 8$) than when they had not heard the novice practice ($M = 47$ ms, $SD = 9$). Accordingly, a 2×2 repeated measures ANOVA showed a significant effect of Familiarity with Novice Performance, $F(1, 11) = 6.003$, $p = .032$ but no significant main effect of Familiarity with Score, $F(1, 11) = .005$, $p = .946$, and no significant interaction between the two factors, $F(1, 11) = .983$, $p = .343$.

For keystrokes from Difficult Passages, i.e. keystrokes that were preceded by a novice's shift in hand position (Fig. 2, Panel B) experts produced lower asynchronies when they were familiar with the score including the novice's part ($M = 45$ ms, $SD = 9$) than when they were not familiar with the score ($M = 53$ ms, $SD = 13$). Accordingly, the ANOVA showed a significant main effect of Familiarity with Score, $F(1, 11) = 14.918$, $p = .003$, but not of Familiarity with Novice Performance, $F(1, 11) = .187$, $p = .674$. There was no significant interaction between the two factors $F(1, 11) = 1.960$, $p = .189$. Signed asynchronies showed a similar pattern. On average signed asynchronies (expert - novice) were negative in all conditions with an overall mean of $M = -4$ ms, $SD = 59$.

3.2. Variability of asynchrony (SD)

Panels C and D in Fig. 2 show the results for the variability of asynchrony in terms of standard deviation in milliseconds separately for

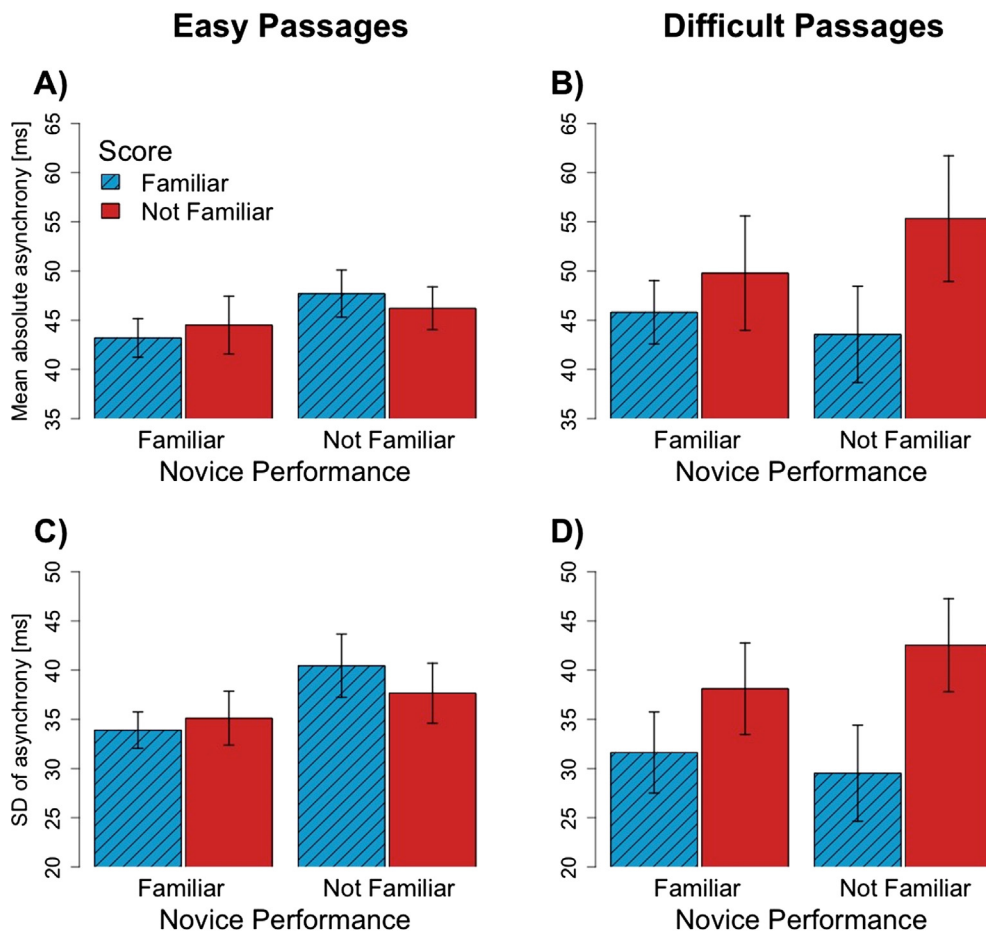


Fig. 2. (A) Mean absolute asynchronies between the keystrokes of experts and novices *not preceded by a hand shift* of the novice. (B) Mean absolute asynchronies between the keystrokes of experts and novices *preceded by a hand shift* of the novice. (C) Variability of asynchronies *not preceded by a hand shift* of the novice. (D) Variability of asynchronies *preceded by a hand shift* of the novice. The error bars represent within-subject confidence intervals according to Cousineau (2005).

Easy and Difficult Passages. The results are in line with the results for mean absolute asynchrony. For keystrokes from Easy Passages (left panel) the standard deviation of asynchrony between expert and novice was lower when experts had heard the novice practice her melody beforehand ($M = 35$ ms, $SD = 6$) than when they had not heard the novice practice ($M = 39$ ms, $SD = 8$). The ANOVA revealed a significant main effect of Familiarity with Novice Performance, $F(1, 11) = 9.297$, $p = .011$ but no significant main effect of Familiarity with Score, $F(1, 11) = .164$, $p = .693$, and no significant interaction between the two factors, $F(1, 11) = 2.809$, $p = .122$.

For keystrokes from Difficult Passages, i.e. keystrokes that were preceded by a novice's shift in hand position (Fig. 2, right panel) experts produced lower variability of asynchronies when they were familiar with the score including the novice's part ($M = 31$ ms, $SD = 8$) than when they were not familiar with the score ($M = 40$ ms, $SD = 11$). Accordingly, the ANOVA showed a significant main effect of Familiarity with Score, $F(1, 11) = 15.809$, $p = .002$, but not of Familiarity with Novice Performance, $F(1, 11) = .237$, $p = .636$. There was no significant interaction between the two factors, $F(1, 11) = 1.199$, $p = .297$.

3.3. Cross-correlation analysis

To provide a further measure of adaptation based on experts' familiarity with a novice's performance, we performed a cross-correlation analysis in which we compared the timing patterns of novices with the corresponding patterns of experts at lag zero. This analysis showed that experts were higher correlated with novices after they had heard the novice performance (mean $R = 0.33$, $SD = 0.20$) than when they had not heard the novice performance (mean $R = 0.21$, $SD = 0.21$). A paired sample t -test showed a significant difference, $t(11) = 2.591$,

$p = .025$. Therefore, the cross-correlation further corroborates the findings obtained in our analysis of absolute asynchronies.

3.4. Random pairing analysis

Furthermore, we assessed whether experts' adaptations were specific to the novice they coordinated with. We compared the correlation of IKI patterns of experts and novices for actual pairs with the correlations for random pairs. Random pairings were constructed in a way so that each expert was paired with each novice except for her actual partner, matched for melody and condition. We then calculated correlations of IKI patterns for the random pairs and computed the upper confidence intervals, a conservative estimate (see Zamm, Pfordresher, & Palmer, 2015). These upper confidence intervals were then compared to the correlation values of the actual pairings. Lag 0 correlations among actual pairs was higher ($M = 0.33$, $SD = 0.22$) than the upper confidence intervals of lag 0 correlations for random pairs ($M = 0.12$, $SD = 0.10$). Using Bonferroni adjusted alpha levels of .025 per test ($.05/2$), the results of paired-samples t -tests showed significant differences between observed correlations and CIs of random-pair-correlations when experts had heard the novice beforehand: $t(11) = 2.98$, $p = .012$ and not when experts had not heard the novice beforehand: $t(11) = 2.16$, $p = .054$.

3.5. Autocorrelation analysis

Finally, to check for indications that novices' variability originates from expressive timing, we analyzed the autocorrelation of timing intervals in novices' performances. We calculated autocorrelation values for each novice and each melody's temporal pattern of inter-keystroke-intervals separately. Based on the melody length of 16 tones, i.e. 15

inter-keystroke-intervals, we were able to compute the correlation coefficients for lags 1 through 11. We found that none of the computed correlation coefficients reached the confidence limits of $\pm 2/\sqrt{n}$, with n being 15 IKI values per melody in our case.

4. Discussion

The present study examined how experts adapt their performance when playing piano duets with novices to ensure joint action coordination despite novices' suboptimal timing. The results showed that there are at least two different factors that enable experts to achieve coordination. During easy passages experts were able to improve coordination when they were familiar with novices' performances. In passages that were difficult for novices, experts were able to improve coordination when they were familiar with novices' scores.

Familiarity with novices' individual performance can only facilitate coordination if experts pick up on timing regularities in novices' performances that reflect novices' idiosyncratic timing patterns. Our results show that experts were able to identify and use such regularities quite well to predict the timing of the novices' performances. Our analysis of random pairings provides evidence that experts indeed identified idiosyncrasies. Furthermore, our finding that the average of the overall signed asynchronies was around zero, i.e., well below the shortest possible reaction time, implies that experts predicted the timing of novice performance (Repp & Su, 2013). Thus, we propose the following mechanism for how experts achieved improved coordination based on being familiar with novices' performances: While listening to the novice, experts compare novices' timing to the timing predicted by their internal models (Repp & Knoblich, 2004). This allows them to generate an error matrix that they can use, during later joint performance, to modulate their predictions of novices' timing, enabling them to reduce asynchronies despite novices' suboptimal timing.

The second way in which experts can improve coordination with a novice is to predict performance difficulties for novices based on symbolic information about their task. In the present study, the crucial information consisted of symbols in the musical score indicating to the expert when the novice would have to change hand positions. Importantly, the marked changes in hand position did not imply any difficulty from an expert's point of view but posed difficulties selectively for novices. Our interpretation that experts relied on symbolic information is supported by the finding that the experts' prior exposure to the novices' sheet music only facilitated joint action coordination in difficult parts that required novices to shift their hand position. The symbols in the musical score might have acted as performance cues (see Ginsborg & King, 2012) that enabled experts to delay their own performance to give novices enough time to change the position of their hand, thereby departing from the constant tempo prescribed in the musical score. To achieve this, experts needed to convert the symbolic information in the musical score into a modulation of the timing of their motor commands.

It is likely that modulations of expert performance due to symbolic cues occur at a longer timescale than modulations of performance based on an error matrix derived from listening to novices' individual performances. Whereas converting symbolic information into a modulation of timing is largely explicit, continuous modulation of one's own predictions based on an error matrix occurs at the level of internal models and is likely to be implicit and to draw on systems that enable massive parallel processing (Wolpert & Kawato, 1998). This interpretation is in line with Ragert et al. (2013) finding that familiarity with playing another's part was detrimental to interpersonal coordination at the key-stroke level associated with shorter timescales, but improved interpersonal coordination of bodily movements associated with longer timescales (see Davidson, 2009) and higher levels of the hierarchical structure of a musical piece (Koelsch, Rohrmeier, Torrecuso, & Jentschke, 2013; Lerdahl & Jackendoff, 1983; Michael & Wolf, 2014).

Despite operating at different timescales, both postulated

mechanisms enable individual musicians to prepare for interpersonal coordination before they actually engage in a joint performance and can contribute to our understanding of the impact of individual and joint rehearsals on joint action coordination (Ginsborg & King, 2012; Ginsborg et al., 2006). These offline-adaptations could reduce the amount of online adaptation and anticipation (Van der Steen & Keller, 2013) required during the joint performance.

A potential alternative explanation for the lower asynchronies in easy passages after experts heard novices perform could be that experts form episodic memories of the novices' performances while listening to them. It has been shown that episodic memory can be used to correctly differentiate between different performances of the same piece with the help of prosodic cues (Palmer, Jungers, & Jusczyk, 2001). However, it is unlikely that episodic memory drove expert performance in the current study. Here, experts were presented with four consecutive performances of a novice and our results suggest that some form of generalization occurred. In the episodic memory account, one would have to assume that each time interval across four highly similar performance was stored, which not only implies a high demand on memory, but would likely result in 'memory mixing' (Van Rijn, 2016). As musicians exhibit strong links between musical action, outcomes and representations thereof (Bangert & Altenmüller, 2003; Brodsky, Kessler, Rubinstein, Ginsborg, & Henik, 2008; Haueisen & Knösche, 2001), an explanation postulating an error matrix based on generalized asynchronies seems more parsimonious and plausible.

Considering our claim that the novices' timing was suboptimal we checked whether there were any indications that novice variability reflected musical intentions through expressive timing. In order to do so, we performed autocorrelation analyses to check for patterns of expressive timing (Desain & de Vos, 1990). We found no significant correlation coefficients, which indicates that expressive timing did not contribute substantially to the novices' variability. Thus, it seems very likely that novices' timing was indeed based on unintended performance difficulties with the task rather than some sort of musical style the novices intended to express.

With regard to implications for teaching, the present study highlights that experts can better ensure successful coordination with novices' suboptimal timing if they have sufficient advance information about novices' task and performance parameters. This may be a precondition for providing coordination scaffolds for novices when joint actions are performed for teaching purposes. The uni-directional feedback in the present study ensured that experts entirely carried the burden of ensuring joint action coordination. They were remarkably flexible in modulating their performance to adapt to novices' suboptimal timing, sacrificing the musical quality of their own performance to achieve interpersonal coordination.

An important goal for future research is to study which additional means experts use to provide a temporal scaffold for novices in teaching situations when novices also receive feedback from experts' performance, and when experts and novices can also visually perceive each other's actions. Some possible options include exaggerating movements (McEllin, Knoblich, & Sebanz, 2017), reducing performance variability (Vesper, van der Wel, Knoblich, & Sebanz, 2011), and relying on emerging leader-follower dynamics (Konvalinka, Vuust, Roepstorff, & Frith, 2010). A further interesting question is whether and when teachers strategically avoid adjusting to novices to maximize learning opportunities for their students. This strategy could be especially useful when a teacher intends to convey particular expressive timing patterns to a student. Generally, studying how teachers adapt to their students' performance and determining when they decide not to adapt seems to be a promising way to further improve our understanding of how procedural skills are learned and taught through social interaction.

Acknowledgements

We thank Dávid Csűrös for his help with data collection. This

research was supported by the European Research Council under the European Union's Seventh Framework Program (FP7/2007-2013)/ERC grant agreement n°609819, SOMICS, and by ERC grant agreement n°616072, JAXPERTISE.

References

- Aoki, T., Furuya, S., & Kinoshita, H. (2005). Finger-tapping ability in male and female pianists and nonmusician controls. *Motor Control*, 9(1), 23–39.
- Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: A longitudinal DC-EEG study. *BMC Neuroscience*, 4(1), 26.
- Brodsky, W., Kessler, Y., Rubinstein, B. S., Ginsborg, J., & Henik, A. (2008). The mental representation of music notation: Notational audiation. *Journal of Experimental Psychology: Human Perception and Performance*, 34(2), 427.
- Butterfill, S. A. (2017). Coordinating joint action. In Kirk Ludwig, & Marija Jankovic (Eds.). *Routledge handbook on collective intentionality*. Routledge handbooks in philosophy. London: Routledge.
- Cousineau, D. (2005). Confidence intervals in within-subject designs. *Tutorials in Quantitative Methods for Psychology*, 1(1), 42–45.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13(4), 148–153.
- Davidson, J. W. (2009). Movement and collaboration in music performance. In S. Hallam, I. Cross, & M. Thaut (Eds.). *The Oxford Handbook of Music Psychology* (pp. 364–376). Oxford: Oxford University Press.
- Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of student's t-test. *International Review of Social Psychology*, 30(1), 92–101.
- Desain, P., & de Vos, S. (1990). Autocorrelation and the study of musical expression. In *Proceedings of the 1990 international computer music conference*. San Francisco: International Computer Music Association.
- Ginsborg, J., Chaffin, R., & Nicholson, G. (2006). Shared performance cues in singing and conducting: A content analysis of talk during practice. *Psychology of Music*, 34(2), 167–194.
- Ginsborg, J., & King, E. (2012). Rehearsal talk: Familiarity and expertise in singer-pianist duos. *Musicae Scientiae*, 16(2), 148–167.
- Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *Neuroimage*, 21(2), 744–750.
- Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience*, 13(6), 786–792. <http://dx.doi.org/10.1162/08989290152541449>.
- Keller, P. E., Knoblich, G., & Repp, B. H. (2007). Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Consciousness and Cognition*, 16(1), 102–111. <http://dx.doi.org/10.1016/j.concog.2005.12.004>.
- Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society B*, 369(1658), 1–12. [20130394](https://doi.org/10.1098/rstb.2013.0394).
- Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: Interactions of perception and action. *Psychological Science*, 12(6), 467–472.
- Koelsch, S., Rohrmeier, M., Torrecuso, R., & Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proceedings of the National Academy of Sciences*, 110(38), 15443–15448.
- Konvalinka, I., Vuust, P., Roepstorff, A., & Frith, C. D. (2010). Follow you, follow me: Continuous mutual prediction and adaptation in joint tapping. *Quarterly Journal of Experimental Psychology*, 63(11), 2220–2230. <http://dx.doi.org/10.1080/17470218.2010.497843> (2006).
- Kourtis, D., Knoblich, G., Woźniak, M., & Sebanz, N. (2014). Attention allocation and task representation during joint action planning. *Journal of Cognitive Neuroscience*, 26(10), 2275–2286.
- Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception: An Interdisciplinary Journal*, 1(2), 229–252.
- Loehr, J. D., & Palmer, C. (2007). Cognitive and biomechanical influences in pianists' finger tapping. *Experimental Brain Research*, 178, 518–528. <http://dx.doi.org/10.1007/s00221-006-0760-8>.
- Loehr, J. D., & Palmer, C. (2011). Temporal coordination between performing musicians. *The Quarterly Journal of Experimental Psychology*, 64(11), 2153–2167. <http://dx.doi.org/10.1080/17470218.2011.603427>.
- 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*. <http://dx.doi.org/10.1037/xhp0000505> Advance online publication.
- Michael, J., & Wolf, T. (2014). Why apply a hierarchical predictive processing framework to music perception and performance? *Empirical Musicology Review*, 9(3–4), 177–182.
- Noy, L., Dekel, E., & Alon, U. (2011). The mirror game as a paradigm for studying the dynamics of two people improvising motion together. *Proceedings of the National Academy of Sciences*, 108(52), 20947–20952.
- Palmer, C., Jungers, M. K., & Jusczyk, P. W. (2001). Episodic memory for musical prosody. *Journal of Memory and Language*, 45(4), 526–545.
- Pecenka, N., & Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Experimental Brain Research*, 211(3–4), 505–515.
- Ragert, M., Schroeder, T., & Keller, P. E. (2013). Knowing too little or too much: The effects of familiarity with a co-performer's part on interpersonal coordination in musical ensembles. *Frontiers in Psychology*, 4(368), 1–15. <http://dx.doi.org/10.3389/fpsyg.2013.00368>.
- Ramenzoni, V. C., Riley, M. A., Davis, T., Shockley, K., & Armstrong, R. (2008). Tuning in to another person's action capabilities: Perceiving maximal jumping-reach height from walking kinematics. *Journal of Experimental Psychology: Human Perception and Performance*, 34(4), 919–928.
- Repp, B. H. (1990). Patterns of expressive timing in performances of a Beethoven minuet by nineteen famous pianists. *The Journal of the Acoustical Society of America*, 88(2), 622–641.
- Repp, B. H., & Knoblich, G. (2004). Perceiving action identity: How pianists recognize their own performances. *Psychological Science*, 15(9), 604–609.
- Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review*, 20(3), 403–452.
- Sebanz, N., & Knoblich, G. (2009). Prediction in joint action: What, when, and where. *Topics in Cognitive Science*, 1(2), 353–367.
- Van der Steen, M. M., & Keller, P. E. (2013). The Adaptation and Anticipation Model (ADAM) of sensorimotor synchronization. *Frontiers in Human Neuroscience*, 7(253), 1–15.
- Van Rijn, H. (2016). Accounting for memory mechanisms in interval timing: A review. *Current Opinion in Behavioral Sciences*, 8, 245–249.
- Vesper, C., van der Wel, R. P. R. D., Knoblich, G., & Sebanz, N. (2011). Making oneself predictable: Reduced temporal variability facilitates joint action coordination. *Experimental Brain Research*, 211(3–4), 517–530. <http://dx.doi.org/10.1007/s00221-011-2706-z>.
- Welsh, T. N., Wong, L., & Chandrasekharan, S. (2013). Factors that affect action possibility judgments: The assumed abilities of other people. *Acta Psychologica*, 143(2), 235–244.
- 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 B: Biological Sciences*, 358(1431), 593–602.
- Wolpert, D. M., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11(7), 1317–1329.
- Zamm, A., Pfordresher, P. Q., & Palmer, C. (2015). Temporal coordination in joint music performance: Effects of endogenous rhythms and auditory feedback. *Experimental Brain Research*, 233(2), 607–615.