Article:

Timmers, R., Macritchie, J., Schabrun, S.M. et al. (3 more authors) (2020) Neural multimodal integration underlying synchronization with a co-performer in music: influences of motor expertise and visual information. Neuroscience Letters, 721. 134803. ISSN 0304-3940

https://doi.org/10.1016/j.neulet.2020.134803

Neural multimodal integration underlying synchronization with a co-performer in music: influences of motor expertise and visual information

Renee Timmers,[a1†] Jennifer Macritchie, [b2] Siobhan M. Schabrun, [c3] Tribikram Thapa, [c4]
Manuel Varlet, [b,d5] Peter E. Keller, [b6]

[a] Department of Music, The University of Sheffield, 34 Leavygreave Road, Sheffield S3 7RD, UK

[b] MARCS Institute for Brain, Behaviour and Development, Western Sydney University, *Locked Bag 1797, Penrith NSW 2751, Australia*

[c] Brain Rehabilitation and Neuroplasticity unit, Western Sydney University, Locked Bag 1797, Penrith NSW 2751, Australia

[d] School of Psychology, Western Sydney University, Locked Bag 1797, Penrith NSW 2751, Australia

[1] r.timmers@sheffield.ac.uk, [2] j.macritchie@westernsydney.edu.au, [3] s.schabrun@neura.edu.au, [4] tribikram.thapa@monash.edu, [5] m.varlet@westernsydney.edu.au, [6] p.keller@westernsydney.edu.au

[†] Corresponding author.

Abstract

Sensorimotor synchronization is a general skill that musicians have developed to the highest levels of performance, including synchronization in timing and articulation. This study investigated neurocognitive processes that enable such high levels of performance, specifically testing the relevance of 1) motor resonance and sharing high levels of motor expertise with the co-performer, and 2) the role of visual information in addition to auditory information. Musicians with varying levels of piano expertise (including non-pianists) performed on a single piano key with their right hand along with recordings of a pianist who performed simple melodies with the left hand, synchronizing timing and articulation. The prerecorded performances were presented as audio-only, audio-video, or audio-animation stimuli. Double pulse Transcranial Magnetic Stimulation (dTMS) was applied to test the contribution of the right dorsal premotor cortex (dPMC), an area implicated in motor resonance with observed (left-hand) actions, and the contribution of the right intraparietal sulcus (IPS), an area known for multisensory binding. Results showed effects of dTMS in the conditions that included visual information. IPS stimulation improved synchronization ability, although this effect was found to reverse for the video condition with higher levels of relevant motor expertise. dPMC stimulation improved or worsened synchronization ability. Level of relevant motor expertise was found to influence this direction in the video condition. These results indicate that high levels of relevant motor expertise are required to beneficially employ visual and motor information of a co-performer for sensorimotor synchronization, which may qualify the effects of dPMC and IPS involvement.

Keywords: synchronization, motor expertise, dorsal premotor cortex, intraparietal sulcus, sensorimotor integration, music performance, visual information.

Highlights:

- Cross-modal binding of visual information may decrease musical synchronization accuracy.
- Abstract and concrete visual movement allow action simulation for synchronization
- IPS and dPMC are causally involved in synchronization if a co-performer is seen and heard
- Instrument-specific expertise improves beneficial use of visual information for synchronization

Introduction

Research on sensorimotor synchronization has uncovered cognitive processes that allow performers to coordinate with high temporal precision. This work highlights the role of allocating attention to self and others (1), predicting the timing of co-performers (2), and reactively correcting for discrepancies in interpersonal timing (3). Notably, successful temporal coordination between performers can be realized on the basis of auditory information only (e.g. 3, 4). Seeing the co-performer in addition to hearing them may under certain circumstances decrease synchronization precision (2). Nevertheless, in natural contexts, performers use both visual and auditory information to guide coordination with co-performers (5; 6). This raises the question of how vision is employed, and how it is integrated with auditory cues and the planning of motor actions to contribute to precise temporal coordination.

The visual channel provides a rich source of information about body movements, cuing observers about ongoing actions and action intentions (e.g. 7). Previous studies have shown the relevance of the observer's motor repertoire for perceptual sensitivity to visually observed actions (6). Indeed, motor resonance to observed actions may contribute causally to improved temporal synchronization in particular when the performer has practiced the co-performers' music (8, 9).

Evidence for the beneficial role of motor resonance for synchronization comes from studies that employed double pulse Transcranial Magnetic Stimulation (dTMS) to temporarily inhibit the involvement of brain areas related to simulating others' actions. dTMS applied to the primary motor cortex reduced accurate adjustment to a tempo-change in auditory stimuli (9), while dTMS applied to the dorsal premotor cortex (dPMC) reduced synchronization in a turn-taking task that presented visual and auditory information of a co-performer (8). Participants performed their part with the right hand, while the pre-recorded pianists performed with their left hand. Motor areas in the right hemisphere were stimulated to target simulations of left-hand co-performer actions rather than interfering with right hand actions.

The present study aimed to investigate the relevance of visual information for motor simulation, and the influence of instrumental expertise, hypothesizing that both strengthen the role of dPMC. Furthermore, we aimed to examine the role of multisensory binding by including dTMS application to the right intraparietal sulcus (IPS, as in 10 and 11). We hypothesized that cross-modal binding is necessary for visual information to (positively or negatively) influence temporal synchronization and therefore predicted that the application of dTMS to IPS may causally affect synchronization accuracy.

These aims led to a study design that combined three audio-visual conditions – audio-only, audio-video, and audio-animation – with three TMS conditions – Sham, dPMC, and IPS, and one between-participant variable of piano expertise, including non-pianists. The effect of dTMS on synchronization ability was tested for each type of audio-visual stimuli, using three measures of "asynchronization". A basic musical synchronization task was used without tempo changes or turn-taking requirements. The audio-animation condition was included to examine the relevance of full video information for action simulation and cross-modal binding, or the sufficiency of movement cues.

Materials and methods[1]

Participants

Twenty–six musically trained participants took part in the study[2]. Two were left handed, and the others right handed. All had normal hearing and normal or corrected-to-normal vision. TMS safety screening was applied to exclude individuals with a history of epilepsy, neurological or musculoskeletal conditions. Participants were grouped into non-pianists, amateur pianists, semi-professional pianists, and professional pianists based on self-report (Table 1), which was preferred over years of experience to avoid influences of age and include differences in level of engagement and proficiency.

The study received ethical approval from University of Western Sydney Human Research Ethics Committee (H9990). Participants gave informed consent and were free to withdraw at any time. None of the participants wished to do so. They received a small fee for participation.

---

[1] See supplementary material for methodological details.

[2] A sample size of 26 was deemed sufficient on the basis of samples of 10 and 15 participants in closely related studies (9; 8).

Table 1. Participant characteristics according to level of piano expertise based on main instrument and self-defined level of musicianship.

Level of piano expertise	Main instrument	Number of participants (female)	Median age (and range) in years	Median (and range) of years of playing an instrument
(3) Professional	Piano	8 (7)	32 (20-44)	27 (15-38)
(2) Semi-professional	Piano	8 (4)	22.5 (18-37)	16 (10-33)
(1) (Serious) amateur	Piano	5 (3)	21 (18-23, 62*)	15 (8-17, 54*)
(0) Non-pianist	Other than piano	5 (2)	25 (22-41)	20 (10-31)

* One data point for age and years of playing an instrument is an outlier and listed separately. All participants were included in the analyses.

Material

An accomplished pianist (serious amateur) performed the left-hand part of four beginner-level melodies (see Figure 1).

The pianist played on a Yamaha Clavinova. MIDI[3] recordings were made to assess note onset and offset timing. Audio and video recordings were made for presentation to participants.

Video recordings were taken from the side, focusing on the left lower arm and hand. To create videos for the audio-animation condition, a green dot was painted on the pianist's hand, and changes in the position of the dot across video samples were tracked using computer vision techniques in a two-dimensional space. Animations were generated that showed the movement of the green dot on top of a flesh-colored rectangle within a black background (see Figure 2).

---

[3] Musical Instrument Digital Interface (MIDI) instruments record onset and offset timing and key velocity for each performed note.
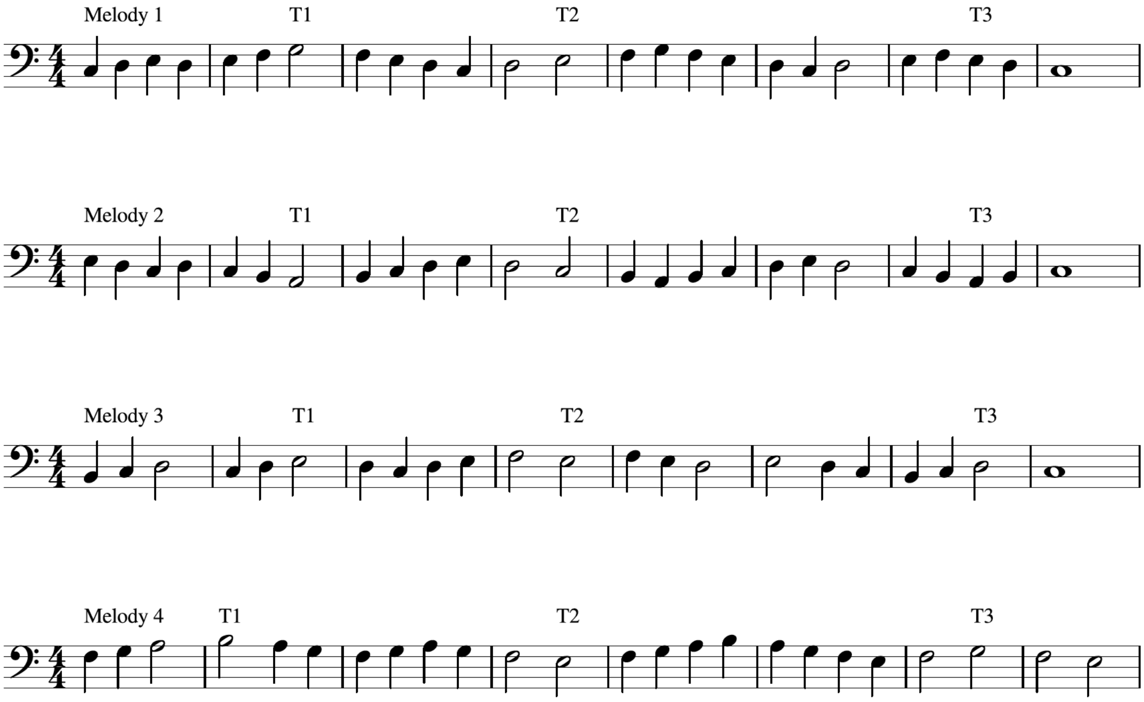
Figure 1. Left hand part of melodies from Bela Bartok's Microcosmos I. Recorded performances of these melodies were used as stimuli. T1-3 indicate target notes for dTMS stimulation.

Figure 2. Illustration of the video and animation stimuli.

The pianist performed the music with some expressive variation in timing and dynamics to create naturalistic stimuli. Three performances were included of each melody. The overall articulation was staccato, the intensity was *forte* (loud) or *mezzo forte* (moderately loud), average tempo ranged between 183-203 BPM and included a modest degree of tempo rubato of on average 7.8% of the average note duration. Half note inter-onset-intervals (IOIs) were on average 624 ms (comparable to IOIs in 8 and 9).

Procedure

Musical preparation

Participants were sent instruction videos at least two weeks before participation to familiarize themselves with the left-hand melodies of the musical stimuli by actively practicing them. Previous studies have shown the benefit of having practiced the other's part for synchronization and motor simulation (8; 9). Video instructions included sound and showed which key to press with what finger at what time in a steady tempo (140 BPM).

Participant's familiarity with the melodies was assessed before participation in the experimental trials. Further practice was given, until the participant performed the melodies without hesitation and errors.

TMS preparation and procedure[4]

Single pulse TMS induced muscle activation in the left hand was measured using electromyography to determine the optimal site for M1 stimulation (hotspot) and resting motor threshold. A standard 70 mm figure-of-eight TMS coil was positioned with the handle pointing postero-laterally at a 45° angle to the sagittal plane. The coil was moved until largest muscle responsiveness was found. Resting motor threshold was subsequently defined as the minimal TMS intensity required to elicit a muscle response of 0.05 mV peak-to-peak amplitude in 5 out of 10 trials.

For the experimental trials, double pulse TMS (dTMS) at 120% resting motor threshold was delivered to P4[5] as a proxy for stimulation of the right IPS (10), to the right dPMC (3 cm anterior to the hot-spot, as in 8), or at 90° coil orientation to Cz for Sham stimulation. This method of localizing dPMC targets a rostral part of dPMC (12).

dTMS was triggered at three score-positions, distributed across two repeats to assure a minimum of 6 seconds between stimulations. In one trial, pulses were triggered at target notes T1 and T3, and, in the other trial, at target note T2 (Figure 1). The first pulse of dTMS was delivered 100 ms before the second, which coincided with the target note onset[6].

---

[4] See supplementary material for further details.

[5] An adjustable EEG cap was used to determine the location of P4.

[6] This was the middle condition in (8), who found no significant difference depending on temporal placement.

Procedure of experimental trials

Participants played along with the presented performances on a silent Yamaha Clavinova with one finger of their right hand. Their task was to synchronize with the "virtual co-performer" as precisely as possible, playing the same rhythm, articulation and dynamics. A computer screen on top of the piano displayed the visual information and audio was presented over speakers. Onset and offset timing of key presses were recorded using MIDI.

A blocked procedure was used in which two melodies were performed in one block and two melodies in the second block. Each block contained all conditions: 2 melodies were presented 6 times (two repetitions of three performances) in each audio-visual condition. This was repeated for each TMS condition. The order of TMS and audio-visual conditions was reversed in the second block of a participant, counter-balancing the orders within participants. Orders were also varied across participants. The total duration of an experimental session was around 90 minutes, including a break between blocks.

Data processing

Data analysis focused on the timing of target notes. Differences between the pre-recorded and participants' performances were measured with respect to the target note's onset timing, duration (note onset to note offset), and the time interval to the next note (IOI). Differences greater than 500 ms were left out, as they were likely the result of an error (e.g. missing note). The distribution of the remaining data was checked for outliers, replacing outliers with the mean plus or minus 2.5 standard deviations. This concerned less than 0.03% for timing, 1.44% for IOI, and 2% for duration difference data. The resulting data set confirmed the assumptions of normally distributed data. Some participants had missing data for short notes. This concerned two participants of the amateur piano group, two participants of the semi-professional piano group and one participant of the professional piano group, leading to reduced degrees of freedom for some of the analyses.

Synchronization measures (sdONSET, sdIOI, sdDUR)

The standard deviations (SD) of the differences in onset-timing (sdONSET), IOI (sdIOI), and duration (sdDUR) were used as summary measures of performance across trials in each condition. The SD captures the variability with which a participant aligned in time with the pre-recorded performance. Separate estimates were made for long and short target notes and for each TMS and AV condition. The standard deviation of timing differences (generally applied to onset or IOI) or "asynchronization" (13) captures the lack of consistency with which performers synchronize. It is in line with well-established timing models that expect asynchrony correction to minimize asynchrony variance (14; 3).

Whilst differences in IOI and onset timing are typically used to investigate synchronization (3; 13), the inclusion of a measure related to duration is uncommon. It was included to reflect the musical task: to truly perform together, musicians align both onset timing and articulation (as discussed in 15).

Data analysis

Analysis of Variance (ANOVA) was used to test the effects of TMS stimulation (TMS) and piano expertise (Piano) on the three dependent variables (sdONSET, sdIOI and sdDUR). The ANOVAs were conducted separately for each type of audio-visual stimuli. Note duration (NDUR, 2 levels, short and long) was included as a within-participants independent variable for analysis related to the musical stimuli.

Results

Table 2 shows the main results of the ANOVA for each dependent variable and each type of audio-visual stimuli. Results will be discussed per audio-visual stimulus type.

Audio-only

For the audio-only conditions, no significant effect of TMS or interactions with TMS were found on any of the dependent variables (see Figure 3 for means and SE per condition). The main effect of NDUR was significant for sdIOI (p<.001) and sdDUR (p<.001), as was the interaction between NDUR and Piano for sdIOI (p=.004), and the main effect of Piano on sdIOI (p=.028). The effects were as expected: asynchronization was larger for long notes than short notes (IOI: M=38.869, SE=1.701 vs. M=24.924, SE=1.625; DUR: M=63.728, SE=2.430 vs. M=29.768, SE=2.009). Furthermore, IOI asynchronization decreased with increasing levels of piano expertise (r=-.480, p=.028, df=20). The interaction between NDUR and Piano was related to a negative correlation between Piano and sdIOI for long notes (r=-.663, p=.001, df=20), but no reliable correlation for short notes (r=.098, p=.673, df=20).

Table 2. Results of the univariate ANOVAs testing the effects of TMS, NDUR and Piano on ONSET, IOI, and DUR for audio-only (top), audio-video (middle) and audio-animation stimuli (bottom). Significant effects are highlighted in bold.

		ONSET		IOI		DUR	
	df[†]	*F*	*r*	*F*	*r*	*F*	*r*
Audio-only							
TMS	2, 38	0.055	.055	0.664	.184	0.366	.138
TMS*Piano	2, 38	0.688	.187	1.412	.263	0.434	.148
NDUR	1, 19	1.964	.307	**38.127*** **	**.817**	**67.374*** **	**.883**
NDUR*Piano	1, 19	1.137	.237	**10.503** **	**.597**	3.322	.386
TMS*NDUR	2, 38	0.071	.063	0.415	.145	0.431	.148
TMS*NDUR*Piano	2, 38	0.107	.077	0.043	.045	2.642	.349
Piano	1, 19	0.243	.114	**5.678***	**.480**	0.101	.071
Audio-video							
TMS	2, 38	0.602	.176	1.657	.283	**4.125***	**.422**
TMS*Piano	2, 38	0.995	.224	1.638	.281	**4.010***	**.417**
NDUR	1, 19	0.020	.032	**28.552*** **	**.775**	**110.091*** **	**.924**
NDUR*Piano	1, 19	0.494	.158	**6.026***	**.491**	2.717	.354
TMS*NDUR	2, 38	0.166	.095	0.252	.114	0.079	.063
TMS*NDUR*Piano	2, 38	1.437	.265	0.124	.084	0.030	.045
Piano	1, 19	**4.435***	**.435**	3.179	.378	1.115	.235
Audio-animation							
TMS	2, 38	**4.092***	**.421**	**4.534***	**.434**	1.755	.291
TMS*Piano	2, 38	2.168	.319	0.571	.170	1.044	.228
NDUR	1, 19	3.020	.370	**23.638*** **	**.744**	**69.889*** **	**.887**
NDUR*Piano	1, 19	0.014	.032	3.006	.370	2.104	.316
TMS*NDUR	2, 38	2.964	.367	1.958	.305	0.230	.110
TMS*NDUR*Piano	2, 38	1.414	.263	0.617	.176	0.395	.141
Piano	1, 19	**4.898***	**.453**	**13.845** **	**.650**	**5.773***	**.483**

[†] degrees of freedom are corrected for violations of sphericity where appropriate

* *p* < .05, ** *p* < .01, *** *p* < .001

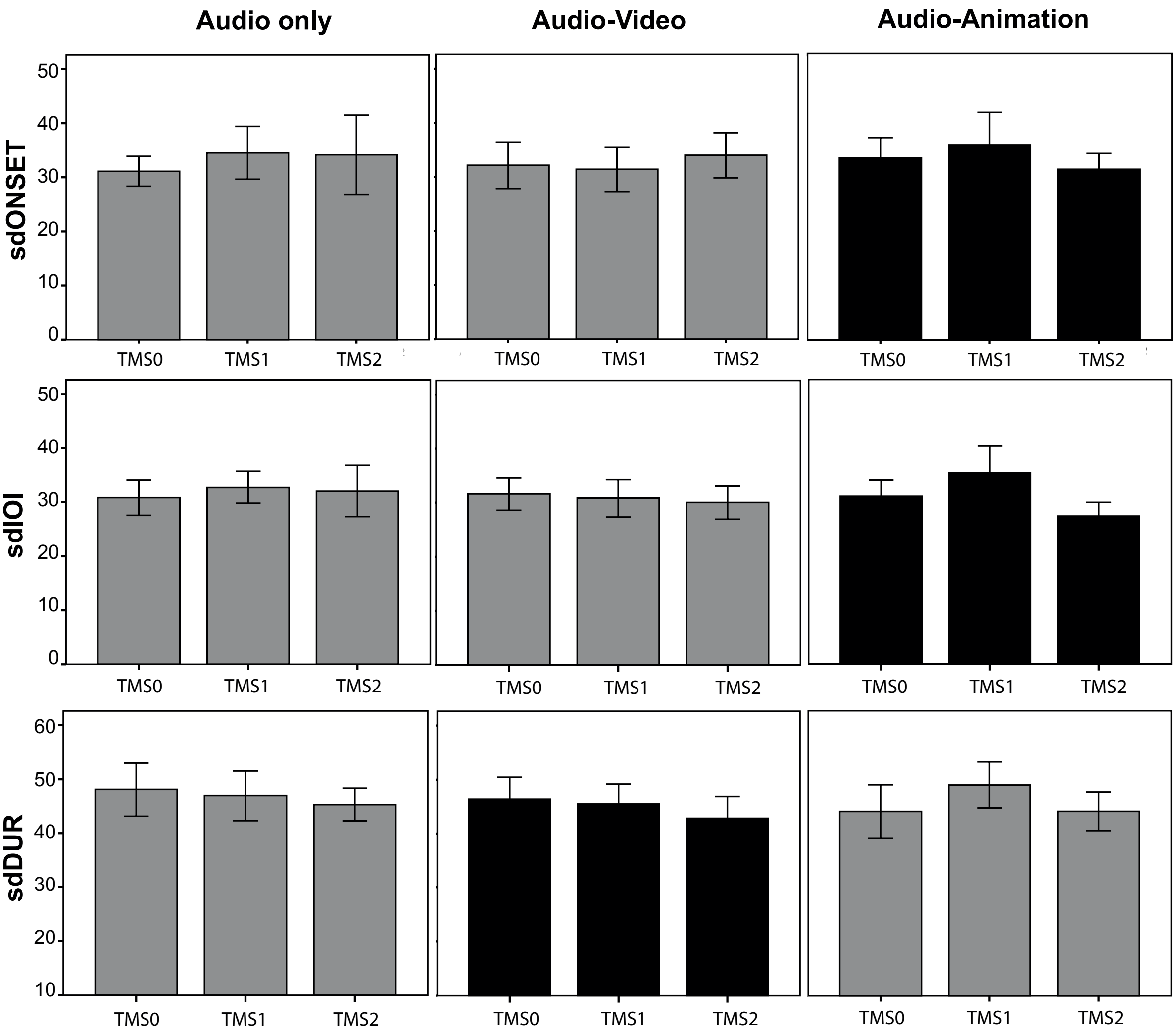

Figure 3. Mean and SE of each asynchronization measure by TMS and audio-visual condition. Significant effects of TMS are in black. Non-significant differences are grey.

Audio-video

For the audio-video conditions, a significant main effect of TMS was found on sdDUR (*p*=.024), and an interaction between TMS and Piano for sdDUR (*p*=.026). This main effect of TMS is illustrated in Figure 3: DUR asynchronization was smallest under IPS stimulation and highest under Sham stimulation. Planned contrasts confirmed significant differences between Sham and dPMC stimulation (*p*=.044), and between Sham and IPS stimulation (*p*=.014).

To investigate the interaction between TMS and Piano on sdDUR, level of piano expertise of individuals was correlated with differences in sdDUR between TMS conditions, which highlights how the effect of TMS (differences between conditions) varies with piano expertise. This analysis showed positive correlations with the difference between dPMC and Sham (r=.481, p=.027, df=20) and with the difference between IPS and Sham (r=.436, p=.048, df=20). These positive correlations relate to changes in the effect of dPMC and IPS stimulation with greater levels of piano expertise, from a decrease in DUR asynchronization ($\Delta M<0$, for expertise levels 0 and 1 under dPMC stimulation; and levels 0, 1, and 2 under IPS stimulation) to an increase ($\Delta M>0$, for expertise levels 2 and 3 under dPMC stimulation, and level 3 under IPS stimulation)[7].

The main effect of NDUR was significant for sdIOI (p<.001) and sdDUR (p<.001), the interaction between NDUR and Piano for sdDUR (p=.024), and the main effect of Piano on sdONSET (p=.049). The effects of NDUR and Piano were as previously observed: asynchronization was larger for long notes than short notes (sdIOI: M=37.644, SE=1.910 vs. M=23.999, SE=1.148; sdDUR: M=61.028, SE=2.244 vs. M=28.676, SE=0.976). The interaction between NDUR and Piano was such that a negative association between Piano and sdIOI was found for long notes (r=-.496, p=.022, df=20, but not for short notes (r=.198, p=.389, df=20). A negative association was present between Piano and sdONSET (r=-.435, p=.049, df=20), indicating improved performance with higher levels of piano expertise.

Audio-animation

For the audio-animation conditions, a significant main effect of TMS was found for sdONSET (p=.025) and sdIOI (p=.017). This effect of TMS showed the same pattern for both measures: asynchronization was relatively large under dPMC stimulation and relatively small under IPS stimulation. Planned contrasts indicated that the pair-wise comparisons with Sham stimulation failed to reach significance for sdONSET (p=.268 for dPMC; p=.320 for IPS). For sdIOI, the difference between Sham and IPS was significant (p=.044), but not the difference between Sham and dPMC (p=.052). As can be seen in Figure 3, the main contrast was between asynchronization under dPMC and IPS stimulation.

Main effects were found of NDUR on sdIOI (p<.001) and sdDUR (p<.001), which were again related to larger asynchronization for long notes than short notes (sdIOI: M=38.704, SE=1.677 vs. M=23.769, SE=1.589; sdDUR: M=63.221, SE=2.400 vs. M=28.090, SE=1.739). Main effects of Piano were observed for sdONSET, sdIOI, and sdDUR. These

---

[7] Scatterplots are provided in the supplementary material.

consisted of significant negative correlations between Piano and sdONSET (*r*=-.453, *p*=.039, *df*=20), Piano and sdIOI (*r*=-.649, *p*=.001, *df*=20), and Piano and sdDUR (*r*=-.483, *p*=.027, *df*=20).

Testing absolute differences in timing and duration

Previous research that is closely related to this study (8) examined synchronization accuracy by measuring the absolute difference in onset timing rather than taking the standard deviation of these differences. The benefit of using asynchronization based on the standard deviation is that it measures accuracy irrespective of the general tendency of a performer to anticipate or lag, or to play more or less legato. Nevertheless, to examine the generalizability of our results to this measure of asynchrony, the analyses were repeated for the absolute differences in onset-timing, IOI and duration, which were log-transformed to correct for positive skew. These analyses showed a significant interaction between TMS and Piano in the video conditions for IOI (*F*(2,44)=3.872, *p*=.028, *r*=.387). This interaction was examined by correlating the differences between TMS conditions with piano expertise, showing a positive correlation for differences in means under IPS compared to Sham stimulation (*r*=.466; *p*=.022, *df*=23). With greater piano expertise, stimulation of IPS led to larger asynchrony ($\Delta M>0$ for levels 1, 2 and 3) rather than smaller ($\Delta M<0$ for level 0). The positive correlation with differences in means under dPMC and Sham stimulation was not significant (*r*=.346; *p*=.098, *df*=23)[8].

Discussion

The asynchronization measures provided converging evidence for a significant effect of dTMS application to dPMC and IPS compared to Sham on musical synchronization, specifically if visual information was present. Furthermore, there was evidence for the effect of dTMS to vary with piano expertise.

The effect of dTMS application to dPMC for the audio-animation condition was as previously found: interference with motor simulation reduces the ability to precisely synchronize with a recorded co-performer (8). In the audio-video condition, stimulation of dPMC was in contrast found to improve synchronization (decrease in sdDUR). This effect

---

[8] Scatterplots are provided in the supplementary material.

was however dependent on level of piano-expertise: with greater expertise the effect reversed and dPMC stimulation increased DUR asynchronization.

The main effect of dTMS application to IPS was a reduction in ONSET and IOI asynchronization in the audio-animation condition. However, within the audio-video condition, this effect interacted with piano expertise for DUR asynchronization, which was also found for IOI in the audio-video condition for the alternative synchronization measure. With greater piano expertise, an increase in asynchronization was observed rather than a decrease.

The main effect of a reduction of asynchronization in the context of IPS and dPMC stimulation requires further explanation. We interpret the reduction as related to an increased complexity of the task with the addition of visual information and motor resonance, which leads to an increase in timing variation. The interaction with piano expertise shows that beneficial employment of visual information for synchronization requires high levels of relevant motor expertise, as also indicated by the multiple significant correlations with expertise for conditions including visual stimuli. Effective use of visual information may involve its use as a source for action simulation (9), and the inclusion of the co-performer's action goals, as in other forms of joint action (16). Notably, the animation condition seemed to be a source for action simulation as much as for visual cuing, given the significant effect of dPMC stimulation, indicating that reduced animations of biological motion provide rich sources of information that can be comparable with full video presentations (17). Future research may reduce the salience of the auditory cues for synchronization, which would allow for more specific testing of the ability to rely on visual information. Furthermore, it will be important to replicate the investigation with a balanced group of expert pianists and non-pianists, as the non-pianist group was small in our sample.

The three asynchronization measures showed varied but not contradictory results. We interpret this finding as an indication that similar processes shape sensorimotor synchronization in terms of onset timing, tempo (IOI), and articulation. It will be of interest to investigate this systematically by isolating instructions to either include onset timing, tempo variation or articulation, thereby controlling the focus of participants' attention to specific performance aspects, which may also reduce inter- and intra-individual variation in the data.

To conclude, the results of this study are consistent with a causal role of both the intraparietal sulcus and the dorsal premotor cortex in synchronization with a musical co-performer. They further indicate a change with expertise in the neurocognitive processes involved in interpersonal synchronization, with greater relevant expertise enhancing the ability to beneficially employ visual and motor information generated by a co-performer.

Acknowledgements

This research was supported by a Leverhulme International Academic Fellowship to RT (IAF-2015-013) and a Future Fellowship grant, Australian Research Council to PK (FT140101162).

Declarations of interest: none.

Credit author statement: RT was responsible for design, data collection, analysis, and write up. JM, MV and PK advised on the design of the study, experimental procedures and manner of data collection and analysis. JM was responsible for stimuli presentation and data recording, including creating the animations. SS and TT advised on and assisted with TMS procedures. TT assisted with data collection. All co-authors provided feedback and input on revisions of the manuscript.

References

1) P.E. Keller, Attentional resource allocation in musical ensemble performance, Psychol. Music 29 (2001) 20-38.
2) P.E. Keller, M. Appel. Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles, Music Percept. 28 (2010) 27-46.
3) B.H. Repp, Sensorimotor synchronization: a review of the tapping literature, Psychon. Bull. Rev. 12 (2005) 969-992.
4) W. Goebl, C. Palmer, Synchronization of timing and motion among performing musicians, Music Percept. 26 (2009) 427-438.
5) L. Bishop, W. Goebl, When they listen and when they watch: Pianists' use of nonverbal audio and visual cues during duet performance, Musicae Scientiae 19 (2015) 84-110.
6) C. Wöllner, R. Cañal-Bruland, Keeping an eye on the violinist: motor experts show superior timing consistency in a visual perception task. Psychol. Res. 74 (2010) 579-585.

7) H. Bekkering, S.F. Neggers, Visual search is modulated by action intentions, Psychol. Sci., 13 (2002) 370-374.

8) Hadley, L. V., Novembre, G., Keller, P. E., & Pickering, M. J. (2015). Causal role of motor simulation in turn-taking behavior. *Journal of neuroscience*, *35*(50), 16516-16520.

9) G. Novembre, L.F. Ticini, S. Schütz-Bosbach, P.E. Keller, Motor simulation and the coordination of self and other in real-time joint action, Soc. Cognit. Affect. Neurosci. 9 (2014) 1062-1068.

10) N. Bien, S. ten Oever, R. Goebel, A.T. Sack, The sound of size: crossmodal binding in pitch-size synesthesia: a combined TMS, EEG and psychophysics study, NeuroImage 59 (2012) 663-672.

11) S. Pasalar, T. Ro, M.S. Beauchamp, TMS of posterior parietal cortex disrupts visual tactile multisensory integration, Eur. J. Neurosci. 31 (2010) 1783-1790.

12) H.R. Siebner, S.R. Filipovic, J.B. Rowe, C. Cordivari, W. Gerschlager, J.C. Rothwell, R.S. Frackowiak, K.P. Bhatia, Patients with focal arm dystonia have increased sensitivity to slow-frequency repetitive TMS of the dorsal premotor cortex, Brain 126 (2003) 2710-2725.

13) R.A. Rasch, Timing and synchronization in ensemble performance, in: J.A. Sloboda (Ed.), Generative processes in music: The psychology of performance, improvisation, and composition, Clarendon Press, Oxford, 1988, pp. 70-90.

14) D. Vorberg, H.H. Schulze, Linear phase-correction in synchronization: Predictions, parameter estimation, and simulations, J. Math. Psychol. 46 (2002) 56-87.

15) P.E. Keller, Ensemble performance: Interpersonal alignment of musical expression, in: D. Fabian, R. Timmers, & E. Schubert (Eds.), Expressiveness in music performance: Empirical approaches across styles and cultures, Oxford University Press, Oxford, 2014, pp. 260-282

16) L.M. Sacheli, G. Tieri, S.M. Aglioti, M. Candidi, Transitory inhibition of the left anterior intraparietal sulcus impairs joint actions: A continuous theta-burst stimulation study, J. Cognit. Neurosci. 30 (2018) 737-751.

17) R. Blake, M. Shiffrar, Perception of human motion, Annu. Rev. Psychol. 58 (2007) 47-73.

Supplementary materials and methods

These supplementary materials provide complementary information to the materials and method description provided in the main manuscript. Information about participants, material, musical preparation, synchronization measures and data processing are available in the main document.

Design

Three audio-visual conditions (AV– audio only, audio-video, and audio-animation) were combined with three dTMS conditions (TMS – Sham, dPMC, IPS) and one between-participants regression co-variate of level of piano expertise (Piano, scale from 0-3). Within each of the experimental conditions of dTMS stimulation and audio-visual information, multiple stimuli were included to obtain a robust, average estimate of the dependent synchronization measures. Stimuli consisted of four simple melodies, three performances of each melody (referred to as versions) and two repetitions of each stimulus. Because the melodies that were performed included two different note durations (half notes and quarter notes), which has been shown to influence synchronization precision, note duration (NDUR) was added as a third independent within-participants variable, in addition to AV and TMS.

Animation stimuli

Video recordings were taken from a side angle, with the region of interest encompassing the hand and forearm movements of the pianist (see Figure 2 of main manuscript). These movements are directly related to sound-production and may therefore be particularly useful for note-to-note synchronization and action simulation. The animation stimuli were generated from the video recording and showed the movement of the left hand of the recorded pianist in abstract form. To create the animations, a green dot was painted on the hand (see Figure 2, left panel of main manuscript), and changes in the position of the green dot across video samples were tracked using computer vision techniques (color segmentation and blob tracking) in a two-dimensional space. To create the animation, the x, y coordinates, representing height and movement towards and away from the piano, were rendered into a green dot of similar size to the original stimuli, and placed on top of a flesh-colored rectangle within a black background to simulate the hand movement and to create a similar degree of movement in the animation as in the video.

TMS preparation and procedure

To determine TMS stimulation locations and stimulation thresholds, TMS induced muscle activation in the left hand was measured using electromyography (EMG). Surface EMG was recorded using dual electrodes (Ag-AgCl, Noraxon dual electrodes, product #272S, inter-electrode distance 2.0 cm) placed in a belly-tendon montage on the relaxed left Abductor pollicis brevis (APB) and first dorsal interosseous (FDI) muscles. The ground electrode was positioned on the left olecranon and EMG signals were amplified and digitized at 4 kHz using an ADInstruments Dual Bio amplifier and PowerLab 16/30 recording system (ADInstruments Pty Ltd., Australia).

TMS was delivered using a standard 70 mm figure-of-eight coil connected to a magnetic stimulator (Magstim 200, Magstim Co. Ltd. Dyfed, UK). Single pulse TMS was used to determine the optimal site for right M1 stimulation (termed "hot-spot") and resting motor threshold. The coil was positioned with the handle pointing postero-laterally at a 45° angle to the sagittal plane. The coil was moved until largest muscle responsiveness was found. Subsequently, resting motor threshold was defined as the minimal TMS intensity required to elicit a muscle response of 0.05 mV peak-to-peak amplitude in 5 out of 10 trials from the relaxed FDI muscle.

For the experimental trials, double pulse TMS (dTMS) was employed at 120% resting motor threshold. Average stimulation level was 55.75 (SD=14.28). dTMS was delivered to P4[9] as a proxy for stimulation of the right intraparietal sulcus (1). The dorsal premotor area was targeted by moving the coil 3 cm anterior from hot-spot, keeping the coil orientation constant (as in 2). Sham stimulation was obtained by tilting the coil 90° away from the scalp, while positioning the wings of the coil to Cz[10] and the handle pointing backwards. These three stimulation methods made up the dTMS stimulation conditions: intraparietal sulcus (IPS), dorsal premotor cortex (dPMC) and sham. It should be noted that the method to localize the IPS was less precise than the manner of localizing the dPMC, which was unavoidable in the context of this study. Note as well that this method of localizing dPMC targets a rostral part of dPMC (see 3).

---

[9] An EEG cap adjusted to the size of the participant's head was used to determine the location of P4.

[10] Cz was determined as the intersection between tape measures from the nasion to the inion and from the left to the right ear.

In accordance with the procedure of Hadley et al. (4), the timing of the dTMS was synchronized with the onset of notes in the audio-visual stimuli: the first pulse was delivered 100 ms before note onset, while the second pulse coincided with a note onset[11]. dTMS was triggered at three score-positions within a melody. These three triggers were distributed across two trials to assure that the time between double pulses was greater than 6 seconds allowing for recovery in between. Consequently, in one trial, pulses were triggered at score location T1 and T3, and, in the other trial, at score location T2 (see Figure 1 of main manuscript). These score locations were towards the end of a sub-phrase, often coinciding with a relatively long note duration (half note). These locations were chosen as local tempo typically changes relatively strongly around group boundaries, which increases demands on temporal coordination.

Procedure of experimental trials

Participants were tested in a spacious lab that was sound attenuated to block out external sounds. The room was well lit, without being overly bright. One or two experimenters were present in the room to facilitate the running of the experiment. Participants were seated at a Yamaha Clavinova Piano with a computer screen on top that displayed the visual information. Participants were seated in a large chair for experimental purposes – with a high back and flat seat. This allowed the experimenter to rest their arm on the back when holding the TMS coil in place. The angle of participants arms was around 90° with the piano keyboard. Participants were asked to play along and synchronize with the virtual co-performer, playing the same rhythm, articulation and dynamics, but using only one key on the piano keyboard. The decision to use one key instead of playing a full melody was taken to minimize differences in difficulty of the task between pianists and non-pianists. Participants pressed this key with a finger of the right hand. For most participants, this was the right index finger, but one participant preferred to perform with the thumb. Key presses did not trigger sounds. However, some auditory feedback was present from the noise of the keypresses themselves. Participants practiced the task before starting the experimental trials. Onset time and offset time of each key press were recorded using MIDI.

---

[11] Hadley et al. (4) compared the effect of three temporal placements of the dTMS pulses. They found no significant difference in effect depending on the temporal placement. A double pulse of 100 ms before note onset and the second coinciding with the note onset was the middle condition.

A blocked procedure was used in which two melodies were performed in one block and two melodies in the second block. Each block contained the three dTMS (referred to as TMS) conditions of Sham, dPMC and IPS stimulation. Each TMS condition contained all three audio-visual conditions. Within a certain audio-visual condition, the two melodies were performed twice in three different recorded versions. The order of TMS conditions, audio-visual conditions and melodies was varied across participants. Furthermore, the order of TMS and audio-visual conditions was reversed in the second block of a participant, allowing for counter-balancing of orders within participants as well as varying across participants. The full experiment consisted of 216 trials in total: 4 melodies × 3 versions × 2 repetitions × 3 TMS conditions × 3 audio-visual conditions. The experimental trials included 216 dTMS pulses and 144 sham dTMS pulses. The total duration of an experimental session was around 90 minutes.

Timing measurements and data processing

The timing of the presentation of stimuli and the triggering of dTMS pulses were controlled using the software program Presentation (Neurobehavioral Systems Inc., Albany, CA, USA) to assure synchronized stimulation. The audio of the stimuli was recorded together with the MIDI performances of the participants using a music sequencing program, to assure synchronized recording of the stimuli and MIDI performances. Participants' MIDI performances were then compared to the original MIDI recordings of the presented stimuli. MIDI was used for timing measurement as it has a high temporal resolution of 1 ms and provides unambiguous timing measurements. Latency in the region of 1-8ms may be introduced by using a USB-MIDI interface to connect the Clavinova to a computer (5). Any such error is commensurate across conditions.

Data analysis focused on the timing of the note coinciding with the second pulse of the dTMS, which we refer to as the target note. Differences between the pre-recorded and the live performance were measured with respect to the target note's onset timing, duration (note onset to note offset), and the inter-onset-interval (IOI) to the next note.

Notes showing differences in onset-timing, IOI or duration greater than 500 ms were omitted from the analysis as they were likely the result of an error (e.g. missing note). The distribution of the remaining data was checked for outliers, as outliers may unduly affect statistical measures and production tasks may generate highly variable data that may not be errors. Data outside the range of the mean plus or minus 2.5 standard deviations were replaced with the respective maximum or minimum value of the mean plus or minus 2.5

standard deviations. This concerned less than 0.03% of the timing difference data, less than 1.44% of the IOI difference data, and less than 2% of the duration difference data. The resulting data set confirmed the assumptions of normally distributed data. Some participants had missing notes. This concerned two participants of the amateur piano group, two participants of the semi-professional piano group and one participant of the professional piano group, leading to reduced degrees of freedom for some of the analyses.

Supplementary results

The main results of the study are reported in the main manuscript. These supplementary results provide scatterplots that illustrate in more detail the interaction effects found and discussed in the main document.

Significant interactions were found in the audio-video condition between the effects of TMS and piano expertise. The scatterplots below are provided to gain further insight into these interactions. Figure 1 shows the relationship between level of piano expertise and differences between TMS conditions, specifically the differences in sdDUR under Sham stimulation and dPMC stimulation (left), and differences in sdDUR under Sham and IPS stimulation (right). Figure 2 shows an alternative representation of the same interaction. It plots the relationship between piano expertise and sdDUR for each TMS condition (Sham, dPMC and IPS). The figures show that the advantage of more expert performers disappears under dPMC and IPS stimulation. More precisely, as shown in Figure 1, this change in effect of piano expertise is related to a difference in the effect of dPMC and IPS stimulation: these may benefit synchronization (decreasing asynchronization) in less expert performers, while they disadvantage synchronization in more expert performers (increasing asynchronization).

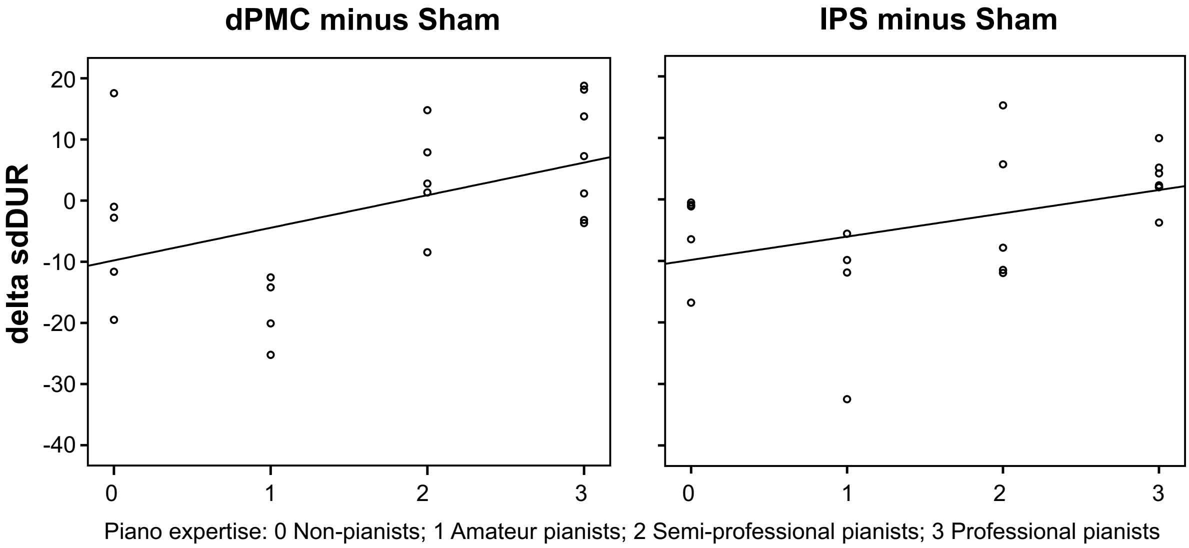

Figure 1. Scatterplots showing the relationship between piano expertise (0-3, non-pianist to professional pianists) and differences in participants' sdDUR under dPMC vs. Sham (left panel), and IPS vs. Sham stimulation (right panel).

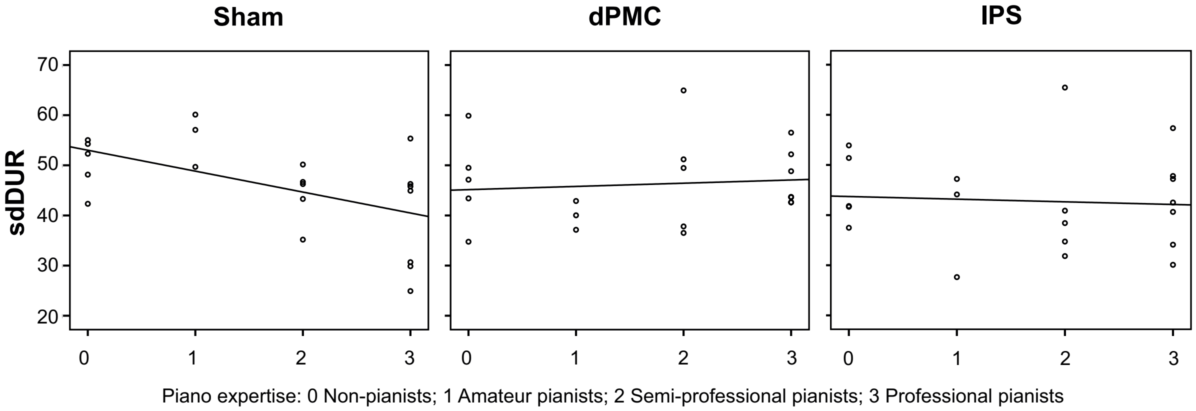

Figure 2. Scatterplots showing the relationship between piano expertise (0-3, non-pianist to professional pianists) and participants' sdDUR in three TMS conditions: Sham, dPMC and IPS stimulation.

Figures 3 and 4 show the parallel relationships with piano expertise for the absolute asynchrony in IOI in the audio-video condition, which was log-transformed to correct for skewness of the data.

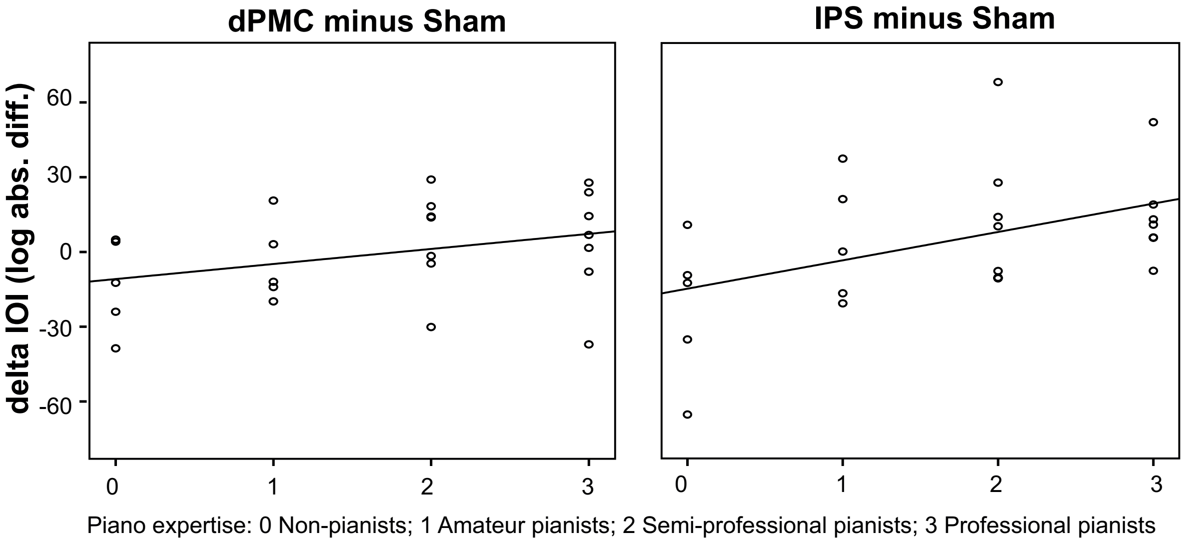

Figure 3. Scatterplots showing the relationship between piano expertise (0-3, non-pianist to professional pianists) and differences in participants' log_absolute_differences in IOI under dPMC vs. Sham (left panel), and IPS vs. Sham stimulation (right panel).

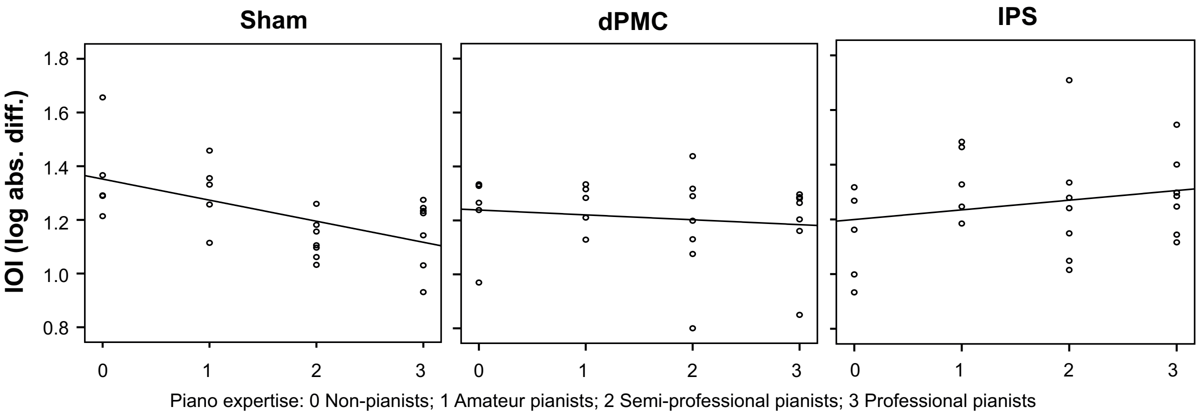

Figure 4. Scatterplots showing the relationship between piano expertise (0-3, non-pianist to professional pianists) and participants' log_absolute_differences in IOI in three TMS conditions: Sham, dPMC and IPS stimulation.

References

1) N. Bien, S. ten Oever, R. Goebel, A.T. Sack, The sound of size: crossmodal binding in pitch-size synesthesia: a combined TMS, EEG and psychophysics study, NeuroImage 59 (2012) 663-672.

2) P.E. Keller, G. Knoblich, B.H. Repp, Pianists duet better when they play with themselves: on the possible role of action simulation in synchronization, Conscious. Cognit. 16 (2007) 102-111.

3) H.R. Siebner, S.R. Filipovic, J.B. Rowe, C. Cordivari, W. Gerschlager, J.C. Rothwell, R.S. Frackowiak, K.P. Bhatia, Patients with focal arm dystonia have increased sensitivity to slow-frequency repetitive TMS of the dorsal premotor cortex, Brain 126 (2003) 2710-2725.

4) L.V. Hadley, G. Novembre, P.E. Keller, M.J. Pickering, Causal role of motor simulation in turn-taking behavior, J. Neurosci. 35 (2015) 16516-16520.

5) B.G. Schultz, The Schultz MIDI Benchmarking Toolbox for MIDI interfaces, percussion pads, and sound cards, Behav. Res. Methods, 51 (2019) 204–234.