Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (ESCOM 2009) Jyväskylä, Finland
Jukka Louhivuori, Tuomas Eerola, Suvi Saarikallio, Tommi Himberg, Päivi-Sisko Eerola (Editors)

The Relationship Between Auditory Imagery and Musical Synchronization Abilities in Musicians

Nadine Pecenka,[*1] Peter E. Keller,[*2]

[*] *Music Cognition and Action Group, Max Planck Institute for Human Cognitive and Brain Sciences, Germany*
[1]`pecenka@cbs.mpg.de`, [2]`keller@cbs.mpg.de`

ABSTRACT

Musical ensemble performance requires precise action coordination. To maintain synchrony in the presence of expressive tempo variations, musicians presumably anticipate the sounds that will be produced by their co-performers and coordinate their own anticipated actions with these predictions. Anticipatory auditory images in pitch and time may facilitate such predictions. Two experiments were conducted to examine the contribution of different aspects of auditory imagery abilities to sensorimotor synchronization (SMS) in musicians. In Experiment 1, the acuity of single-tone pitch images was measured by an adjustment method and by adaptive threshold estimation. Different types of finger tapping tasks were administered to assess SMS. Auditory imagery and SMS abilities were found to be positively correlated with one another and with musical experience. Importantly, however, the imagery/synchronization relationship was only partially mediated by musical experience. In Experiment 2, the acuity of pitch images of short melodic sequences and temporal images of simple rhythmic sequences was assessed by adaptive threshold estimation procedures. SMS was measured by finger tapping. An imagery/sychronization relationship was revealed for temporal imagery ability only. The results of the present experiments are consistent with the notion that auditory imagery ability is of importance for musical synchronization. Furthermore, temporal imagery acuity appears to be more closely related to performance in simple SMS tasks than pitch imagery acuity.

INTRODUCTION

Auditory imagery plays a major role in musical activities (Halpern & Zatorre, 2005). Musicians often hear with their “mind’s ear” when they read musical notation, memorize or compose new music, and they rely on musical images to guide their performances (e.g., Trusheim, 1991; Deutsch & Pierce, 1992). It has recently been suggested that the ability to form auditory images is important for interpersonal action coordination in musical ensembles (Keller, 2008).

To date, only a handful of empirical studies have investigated auditory imagery abilities in musicians. Compared with non-musicians, musically trained individuals have been found to perform better on musical as well as non-musical auditory imagery tasks, but not on a visual imagery task (Aleman, Nieuwenstein, Boecker, & de Haan, 2000). In addition, musicians’ auditory images of single-tone pitches stored in working-memory are more resistant to verbal and visual interference than non-musicians’ images (Pechmann & Mohr, 1992). A recent MEG study showed that imagery of short melodies was strong enough to evoke an imagery mismatch negativity in response to an incorrect external continuation of the melody only in musicians (Herholz et al., 2008). Studies of musicians with different amounts of musical training have found that auditory imagery abilities improve with increasing musical experience. For example, Janata and Paroo (2006) reported a close relationship between the duration of musical training and the acuity of auditory pitch images. Schendel and Palmer (2007) found that more experienced musicians were less disturbed by musical suppression (singing “la”) in a task that required the translation of musical notation into an auditory form.

Whether improvements in auditory imagery ability benefit actual musical performance has received relatively little empirical attention. One relevant study found a positive correlation between pianists’ auditory imagery abilities and success at learning novel piano pieces in the absence of auditory feedback (Highben & Palmer, 2004). Another study revealed a relationship between auditory imagery abilities and interpersonal coordination during duet piano performance (Keller, 2008).

The aim of the present study is to investigate the contribution of auditory imagery abilities to basic sensorimotor synchronization (SMS) processes in musicians. We assume that musicians anticipate future sounds that will be produced by their co-performers and coordinate their own upcoming actions based on these predictions. If such predictions rely on active auditory imagery processes, a positive correlation should be observed between auditory imagery and synchronization skills. We tested this hypothesis in two experiments with a task battery assessing different basic auditory imagery and SMS abilities.

EXPERIMENT 1

Experiment 1 was conducted to examine the relationship between pitch imagery and SMS ability. To assess auditory imagery ability, the acuity of single-tone pitch images was measured by (A1) an adjustment method and (A2) by adaptive threshold estimation. SMS ability was assessed with three finger tapping tasks: (S1) On-beat tapping with a stable metronome, (S2) On-beat tapping with a tempo changing pacing signal, and (S3) Off-beat tapping in antiphase with a stable metronome.

A. Methods

1) Participants. Twenty musicians with varying degrees of musical experience were tested (years playing summed over all instruments: *range* = 5-66, M = 26.8, SD = 14.1). Their mean age was 24.5 years (SD = 2.7). Participants were recruited from a database of participants that took part in previous studies at the Max Planck Institute for Human Cognitive and Brain Sciences. They were systematically recruited to display a large variability in their degrees of musical experience. None of the participants reported possessing absolute pitch.

2) General Procedure. Participants were seated in a laboratory room in front of a computer monitor. Stimuli were presented through Sennheiser headphones at a comfortable sound level. Stimulus presentation and response recording was controlled using MAX/MSP running on a Windows computer. Prior to the commencement of each task completion, participants received brief oral instructions and more detailed on-screen instructions, and they completed a short block of practice trials to ensure that they felt familiar with the task. Upon completing all the tasks, participants filled out a questionnaire addressing their degree of musical experience.

3) Pitch Imagery Tasks. In the two auditory pitch imagery tasks, participants were required to maintain the image of a target tone over a 10-s silent interval and subsequently (A1) to adjust the pitch of a probe tone to match the pitch of the imagined target tone, or (A2) to compare the pitch of a probe tone to the target pitch (sharp vs. flat). In task A1, frequencies were adjusted by moving the bar of a slide control with the computer mouse. Responses in task A2 were given by using the up (sharp) and down (flat) arrow keys of the computer keyboard. No feedback regarding response accuracy was given. Target tones were synthesized in MAX/MSP (www.cycling74.com). These tones were comprised of a fundamental frequency and the next five higher harmonics. Targets with the three base frequencies C_4 (261.36 Hz), G_4 (392 Hz), and C_5 (523.25 Hz) were presented in random order. In task A1, probe tone start frequencies were chosen randomly within a range of 200-1200 Hz. Probe tone duration was variable, depending on the individuals' adjustment times (*range* = 2-40 s). Each of the three frequencies was adjusted 10 times. In task A2, probe tone frequency initially deviated from target tone frequency by 4.5%. Employing a threshold estimation procedure (a two-alternative forced-choice task using a weighted up-down method) to estimate pitch image acuity, the amount of deviation was reduced after each correct response and increased after each error. The number of trials in task A2 was therefore dependent on how soon each individual's discrimination threshold was reached. In both imagery tasks, a 1-s burst of white noise was presented between trials. Participants were explicitly instructed to refrain from making any vocalizations that could help them maintain the pitch of the probe tones.

4) Sensorimotor Synchronization Tasks. SMS ability was assessed with three finger tapping tasks: (S1) On-beat tapping with a stable metronome (500 ms inter-onset interval, IOI), (S2) On-beat tapping with a tempo changing pacing signal (400-600 ms IOI range), and (S3) Off-beat tapping in antiphase with a stable metronome (500 ms IOI). Tasks S1 and S3 consisted of 10 trials with 40 metronome beats. Task S2 comprised 12 trials with 88 beats in which tempo transitions followed sigmoidal functions resembling tempo changes found in music (accelerando and ritardando). Metronome beats were articulated by a sampled bell sound. Finger-tapping performance was recorded using a MIDI percussion pad (Roland SPD-6). No feedback sounds were delivered.

B. Results

1) Auditory Imagery. In the adjustment task (A1), auditory imagery acuity was assessed by computing the mean adjusted difference between target and probe tone frequencies. Probe tone settings differed markedly across individuals, with the majority of individuals' images being mistuned upward (see Table 1). In the adaptive threshold estimation task (A2), auditory imagery acuity was assessed by calculating the just noticeable difference (75% correct threshold) for target-probe frequency discrimination (see Table 1).

Pitch images were less accurate in the adjustment task compared to the adaptive threshold estimation tasks, $t(19) = -5.49, p < .001$, perhaps due to interference caused by the adjustment procedure itself. The two tasks were only moderately correlated [$r(18) = .42, p = .06$], suggesting that they measure different aspects of imagery ability. While both tasks measure the acuity of auditory images, the adjustment task additionally assesses the susceptibility to interference by sounds associated with probe tone adjustments. For each musician, a composite score representing the combination of these different aspects of auditory imagery was computed (by averaging z-transformed single scores). Significant correlations were found with aggregated instrumental experience but not with current amount of practice (see Table 2).

Table 1. Summary statistics for the two auditory imagery and three SMS tasks in Experiment 1

		Mean	SD	Range
Auditory imagery tasks				
A1	Pitch deviation	187.43	148.32	14-440
A2	Pitch deviation	22.66	27.32	4-85
SMS tasks				
S1	MAA	26	15	15-62
	VA	449	254	175-1362
S2	MAA	37	15	21-78
	VA	1744	1441	699-6332
	Prediction/tracking	1.14	.20	.82-1.55
S3	MAA	51	41	10-132
	VA	4794	6241	141-18017

Notes. MAA = Mean absolute asynchrony. VA = Variance of asynchronies. Units in auditory imagery tasks are in cent. A deviation of 100 cent corresponds to one semitone on a Western musical scale. Units in the SMS tasks are in ms.

Table 2. Correlations between musical experience and measures of auditory imagery and SMS in Experiment 1

	Auditory imagery			SMS	
	A1	A2	Comp. score	MAA	VA
Instrument play[1]	-.52*	-.54*	-.63**	-.50*	-.51*
Current practice[2]	-.08	-.18	-.15	-.33	-.30

Notes. * $p < .05$; ** $p < .01$. *Df* = 18 for all analyses. MAA = Mean absolute asynchrony. VA = Variance of asynchronies. Comp. = Composite. [1] years of instrument play summed over all instruments [2] hours per week

2) Sensorimotor Synchronization. To estimate SMS ability, mean absolute asynchronies (i.e., the absolute time difference between each metronome beat and the corresponding finger tap) and variance of asynchronies (i.e., the variability of signed within-trial asynchronies) were computed. The three tapping tasks differed significantly on these two measures: $F(2,18) = 12.39$, $p < .001$ for mean absolute asynchrony; $F(2,18) = 9.91$, $p < .01$ for variance of asynchronies. SMS decreased in precision and increased in variability from task S1 through S2 to S3 (see Table 1). Performance was positively correlated between the tasks on both measures and therefore composite scores were computed (by averaging z-transformed single scores from the three tasks). Correlations between SMS and musical experience were qualitatively similar to those found for auditory imagery (see Table 2).

Task S2 allowed us to analyze the degree to which individuals were predicting vs. tracking tempo changes in the pacing signal. Prediction and tracking indices were computed based on lag-0 and lag-1 cross-correlations between inter-tap intervals and metronome IOIs (with results normalized by subtracting the lag-1 autocorrelation of the latter; for details see Repp, 2002). The lag-1/lag-0 ratio reflects whether individuals were predicting (ratio > 1) or tracking (ratio < 1) ongoing tempo changes. The mean of observed ratios was 1.14, indicating that the majority of individuals predicted tempo changes. Prediction/tracking ratios did not correlate with musical experience. However, participants who engaged in more prediction tapped more precisely in tasks S1 [$r(18) = .51$, $p < .05$] and S2 [$r(18) = .66$, $p < .01$].

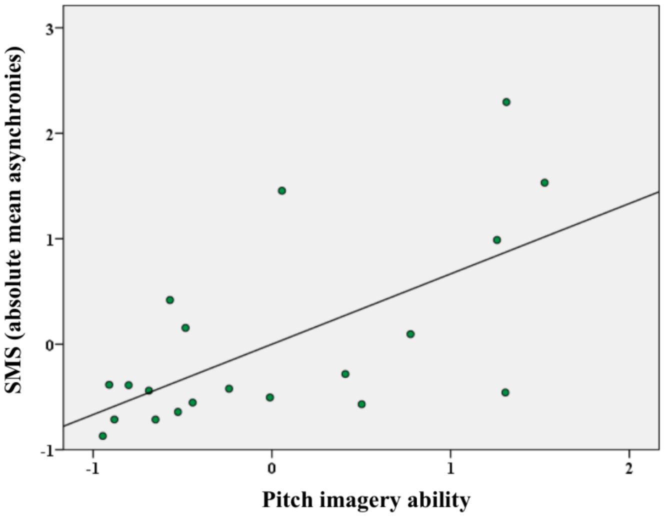

Figure 1. Scatterplot of auditory pitch imagery and SMS composite scores in Experiment 1

3) Relationships Between Auditory Imagery and SMS. Individuals who formed more accurate single-tone pitch images tapped with greater precision [$r(18) = .63$, $p < .01$] and less variability [$r(18) = .62$, $p < .01$] (see Figure 1). Importantly, this relationship between auditory imagery and synchronization ability was not completely mediated by musical experience. When controlling for experience (years of instrument playing), the correlation is reduced but remains significant for mean absolute asynchrony [$r(18) = .48$, $p < .05$] and for variance of asynchronies [$r(18) = .46$, $p < .05$].

Examination of SMS precision in the three tapping tasks separately revealed that correlations with auditory imagery were highest for the more prediction-demanding tapping task: task S2 [$r(18) = .67$, $p < .01$] vs. tasks S3 [$r(18) = .56$, $p < .05$] and S1 [$r(18) = .46$, $p < .05$], although these differences in correlation strength were not significant. Furthermore, prediction/tracking ratios were negatively correlated with auditory imagery composite scores [$r(18) = -.46$, $p < .05$], i.e., individuals who predicted rather than tracked tempo changes in the tapping task (S2) also formed relatively more accurate single-tone pitch images. The above relationships hold when the two auditory imagery tasks are examined separately, although the correlations are more robust for task A2 than for task A1.

C. Discussion

The results of the present study add to a growing body of work showing that auditory imagery ability improves with increasing musical experience. Moreover, we could show that individuals who perform well on auditory pitch imagery tasks are more precise and less variable than others when tapping in synchrony with stable and tempo varying metronomes. Importantly, this relationship between auditory imagery and SMS ability was only partially mediated by musical experience. Further evidence for a link between auditory imagery and anticipatory processes comes from our finding that individuals with relatively good imagery abilities engaged in more prediction when tapping in time with tempo-changing pacing signals.

The relationships observed in the current experiment may seem surprising given that pitch information does not play a major role in the finger tapping tasks administered, where mainly temporal information (the IOIs of metronome beats) was manipulated. Experiment 2 followed up this issue by including temporal imagery tasks.

EXPERIMENT 2

Experiment 2 compared the relative roles of pitch and temporal imagery in SMS. Auditory imagery ability was assessed with one task focusing on pitch imagery acuity (PA) and two tasks examining temporal imagery acuity for constant (TA1) and tempo-changing sequences (TA2).

The pitch imagery task was different from that employed in Experiment 1. In the former experiment, the to-be-imagined tone was heard once and had to be held active in working memory over a long silent interval. To exclude any possible contribution of sensory memory traces in this task in the current experiment, we assessed imagery acuity of tones that were not presented before, i.e. the tones had to be conjured up in working memory. Furthermore, we decided to measure imagery acuity exclusively by an adaptive threshold estimation procedure (two-alternative forced-choice task employing a weighted up-down method). This procedure allows us to estimate each participant's just noticeable difference (75% correct answers) regarding the deviation of the image from perfect tuning or timing more precisely. In contrast to Experiment 1, thresholds were estimated simultaneously for both deviation directions along the pitch and temporal dimension.

SMS ability was assessed with two finger tapping tasks: (S1) On-beat tapping with a stable metronome and (S2) On-beat tapping with a tempo changing pacing signal. Off-beat tapping ability—which was found to be strongly correlated with on-beat tapping ability in Experiment 1—was not assessed in the current experiment.

A. Methods

1) Participants. Forty-seven students from a class of systematic musicology at the University of Leipzig were tested in return for partial course credit. Their mean age was 24.3 years (*SD* = 2.5). Participants varied in their degree of musical experience (years of playing summed over all instruments: *range* = 4-41, *M* = 21.4, *SD* = 8.8). In contrast to Experiment 1 no professional musicians were included in this sample. None of the participants reported possessing absolute pitch.

2) General Procedure. The general procedure and apparatus were identical with those employed in Experiment 1. Participants received detailed instructions and a short block of practice trials before completing the task. At the end of the session, a questionnaire addressing their degree of musical experience was administered.

3) Pitch Imagery Task. In the pitch imagery task (PA), participants heard three tones of an ascending diatonic scale and had to imagine its continuation for two consecutive tones. The mental image of the second continuation tone then had to be compared in pitch (sharp vs. flat) to a presented probe tone. Responses were given by using the up (sharp) and down (flat) arrow keys of the computer keyboard. No feedback regarding response accuracy was given. Presented sounds were synthesized in MAX/MSP. These tones included a fundamental frequency and the next seven higher harmonics, and were presented with envelopes to resemble a piano sound (10 ms attack and 300 ms decay time). The standard inter-onset interval (IOI) of the first three tones was 600 ms. The probe tone was presented 2400 ms after the onset of the third presented tone. The to-be-imagined target tone frequencies were drawn randomly from a range of C_4 (261.36 Hz) to C_5 (523.25 Hz). Probe tone frequencies initially deviated from target tone frequencies by 100 cent. Task duration was dependent on how soon each individual's discrimination threshold was reached. Participants were instructed to refrain from making any vocalizations that could help them maintain the pitch of the probe tones.

4) Temporal Imagery Tasks. In the two temporal imagery tasks (TA1 and TA2), participants were required to imagine the continuation of a presented five-beat sequence for two consecutive beats and judge the temporal acuity (too early vs. too late) of a consecutively presented probe beat. Responses were given with the back (early) and forward (late) arrow keys of the computer keyboard. No feedback regarding response accuracy was provided. Metronome beats were sampled bell sounds produced by a MIDI percussion pad (Roland SPD-S). The presented beat sequences were either of constant tempo (TA1; 400 or 500 ms IOI) or included a tempo change (TA2; 400-500 ms or 500-400 ms IOI) that was designed to resemble tempo changes used for expressive means in performed music (accelerando and ritardando). The initial time deviation of the probe beat was set to 25% of the target IOI. Task duration in both tasks depended on how soon each individual's discrimination threshold was reached. Participants were instructed to refrain from making any vocalizations or movements that could help them keep time in the context of the beat sequence.

5) Sensorimotor Synchronization Tasks. Participants were asked to tap a finger in synchrony with (S1) a stable metronome (400 and 500 ms IOI) or (S2) a tempo changing pacing signal (400-500 ms IOI range). Tasks S1 consisted of 5 trials with 40 metronome beats for each tempo condition. Task S2 comprised 10 trials with 78 beats in which tempo transitions followed sigmoidal functions resembling tempo changes found in music. Stimuli sounds and apparatus were identical with those employed in Experiment 1.

B. Results

1) Auditory Imagery. In the pitch imagery task (PA), auditory pitch imagery acuity was assessed by calculating the just noticeable difference (75% correct threshold) for target-probe frequency discrimination. In the temporal imagery tasks (TA1 and TA2), the just noticeable difference for temporal deviations of the probe beat was computed. In contrast to Experiment 1, separate thresholds for both directions of deviation along pitch (PA) and time (TA1 and TA2) dimensions were estimated. The deviation of each threshold from 0 (maximum accuracy) is used to describe the wideness of the image. To yield an overall measure of image sharpness for each task, the difference between the two threshold estimates was calculated. Descriptive statistics for both tuning and timing thresholds are given in Table 3.

Pitch imagery acuity differed markedly across individuals. Over all participants an average tendency for pitch images to be wider in the negative (flat) direction was found, *t*(46) = -1.86, *p* = .069, i.e., participants' images tended to be tuned slightly downwards in pitch. Similarly, inter-individual variability was high in the temporal imagery tasks (TA1 and TA2). When the continuation of a constant tempo sequence had to be imagined (TA1), a tendency for images to be wider in the positive (late) direction was found *t*(46) = -2.01, *p* = .051, i.e., participants tended to imagine a slightly slower tempo. Images of beat sequences that included tempo changes (TA2) showed biases that were dependent on the direction of tempo change. When a decelerating sequence had to be continued mentally, images were much wider in the negative direction, i.e., tones were expected to occur earlier. In contrast, for the imagined continuation of an accelerating sequence, the deviation of images was larger in the positive direction, i.e., tones were expected to occur later. In about 30% of the participants this bias was so strong that even on-time probe beats were misjudged as being either too early (acceleration) or too late (deceleration). Due to this unexpected bias, thresholds could not be reliably estimated for these conditions and they were excluded from analyses. Average threshold scores were computed over the two remaining unbiased conditions. Taken together, initially presented tempo changes in task TA2 were consistently underestimated when they had to be continued using imagery. Moreover, the deviation of the temporal image was significantly larger in the deceleration condition, $t(46) = 3.52$, $p < .01$.

Performance on the two temporal imagery tasks (TA1 and TA2) was moderately correlated ($r = .42$; $p < .01$) and therefore a composite score for temporal imagery ability was computed (by averaging z-transformed single scores). In contrast, no significant relationship between absolute deviation of pitch images and temporal imagery ability (composite) was found, $r(45) = .18$; $p = .23$. Neither pitch imagery nor temporal imagery ability were significantly related to self-report measures of musical experience (see Table 4).

Table 3. Summary statistics for the two auditory imagery and two SMS tasks in Experiment 2

		Mean	SD	Range
Auditory imagery tasks				
PA	**Pitch imagery**			
	Lower threshold	-83.20	42.52	-172-(-3)
	Upper threshold	68.14	35.76	3-145
	Difference	-15.07	55.51	-147-80
TA1	**Temporal imagery – constant tempo**			
	Lower threshold	-88.58	42.80	-201-(-23)
	Upper threshold	110.96	50.55	22-199
	Difference	22.38	76.51	-179-165
TA2	**Temporal imagery – changing tempo**			
	Tempo acceleration			
	Lower threshold	-44.11	44.17	-165-0
	Upper threshold	146.56	39.31	42-195
	Tempo deceleration			
	Lower threshold	-181.26	47.41	-241-(-56)
	Upper threshold	62.75	68.05	0-225
SMS tasks				
S1	MAA	23	10	10-57
	VA	424	210	150-1093
S2	MAA	34	14	17-76
	VA	1281	682	419-3392
	Prediction/tracking	1.02	.06	.87-1.13

Notes. MAA = Mean absolute asynchrony. VA = Variance of asynchronies. Units in auditory imagery tasks are in cent (PA) and ms (TA1 and TA2). Units in the SMS tasks are in ms.

2) *Sensorimotor Synchronization*. As for Experiment 1, mean absolute asynchronies and variance of asynchronies were computed as indicators for SMS ability. The two tapping tasks differed significantly on these two measures: $t(46) = -6.39$, $p < .001$ for mean absolute asynchrony; $t(46) = -10.50$, $p < .001$ for variance of asynchronies. Lower precision and higher variability was found in tapping with a tempo changing pacing signal (S2). Performance on the two tasks was positively correlated on both measures and therefore composite scores were computed (by averaging z-transformed single scores). There were no significant correlations between SMS and musical experience (see Table 4).

Similar to Experiment 1, the degree to which individuals were predicting vs. tracking tempo changes in tapping task S2 was analyzed. The mean of observed ratios was 1.02, which only marginally suggests prediction of the tempo changes in the pacing signal, $t(46) = 1.95$, $p = .06$. However, when looking at the participants' single scores, 30 out of 47 showed prediction/tracking ratios that exceeded the score 1, suggesting that the majority of individuals indeed favoured a prediction strategy. Prediction/tracking ratios did not correlate with musical experience. Nevertheless, participants who engaged in more prediction also tapped more precisely in task S1, $r(46) = -.60$, $p < .001$ both for mean absolute asynchrony and variance of asynchronies.

Table 4. Correlations between musical experience and measures of auditory imagery and SMS in Experiment 2

	Auditory imagery		SMS	
	PA	Comp. score (TA1 + TA2)	MAA	VA
Instrument play$^{1)}$	-.27	-.04	-.20	-.19
Current practice$^{2)}$	.14	.11	-.09	-.08

Notes. *Df* = 45 for all analyses. MAA = Mean absolute asynchrony. VA = Variance of asynchronies. Comp. = Composite. $^{1)}$ years of instrument play summed over all instruments $^{2)}$ hours per week

3) *Relationships Between Auditory Imagery and SMS*. Individuals who formed more accurate temporal images tapped with greater precision [$r(46) = .45$, $p < .01$] and less variability [$r(46) = .51$, $p < .01$]. In contrast, no such relationship was revealed for pitch imagery and SMS ability; $r(46) = .17$, $p = .26$] for mean absolute asynchrony and [$r(46) = .04$, $p = .81$ for variance of asynchronies (see Figure 2). In line with these findings, prediction/tracking ratios were negatively correlated with auditory temporal imagery composite score [$r(45) = -.50$, $p < .001$], but not so with pitch imagery acuity [$r(45) = -.21$, $p = .15$]. This indicates that individuals who predicted rather than tracked tempo changes in the tapping task (S2) formed more accurate images in the temporal but not in the pitch domain.

C. Discussion

The results of the second experiment support the notion that auditory imagery ability is of importance for SMS performance. However, the observed results highlight the relative importance of temporal imagery ability over pitch imagery ability in finger-tapping tasks. While individuals who perform well on auditory temporal imagery tasks are more precise and less variable than others when tapping in synchrony with stable and tempo varying metronomes, no such imagery/synchronization relationship was found for pitch imagery acuity. This pattern is also reflected in the role of anticipation in SMS with tempo-changing pacing signals. Particularly those individuals with relatively good temporal (but not necessarily good pitch) imagery abilities engaged in more prediction during finger-tapping.

To conclude, temporal imagery acuity appears to be more closely related to performance in simple SMS tasks than pitch imagery acuity. Thus, musicians whose auditory temporal images of upcoming sequential sounds are relatively accurate, are also more precise in synchronizing their movements with the actual sounds.

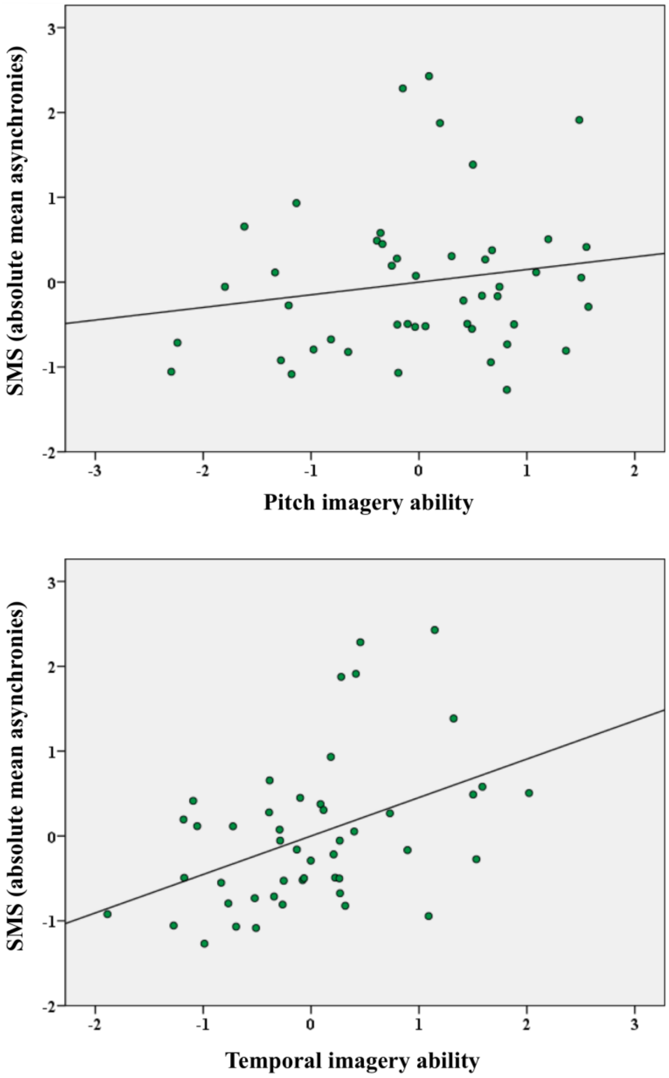

Figure 2. Scatterplot of auditory imagery and SMS composite scores shown separately for pitch imagery (upper panel) and temporal imagery (lower panel).

GENERAL DISCUSSION

The findings of the current study are consistent with the proposal that auditory imagery plays a role in musical synchronization. Evidence for a link between auditory imagery and anticipatory processes comes from our findings that individuals with relatively good pitch (Experiment 1) and temporal (Experiment 2) imagery abilities engaged in more prediction when tapping in time with tempo-changing pacing signals. Unexpectedly, relationships between pitch imagery and SMS ability were not significant in Experiment 2.

Two reasons could mainly account for such discrepancies between experiments: First, different paradigms were used to measure pitch imagery acuity in Experiment 1 and 2. While in Experiment 1 individuals were required to simply maintain the image of a presented tone over a long silent interval, the to-be-imagined tone had to be completely self-created in Experiment 2. Besides the possibility that different cognitive processes might be involved in the two tasks, the pitch imagery task in Experiment 2 was also of higher difficulty, as is reflected in larger pitch imagery thresholds (compare task A2 in Table 1 and task PA in Table 3). Second, sample differences also might have contributed to differences in the results of the two experiments. For Experiment 1, we recruited participants with varying degrees of musical experience, ranging from relative beginners to professional musicians. In contrast, Experiment 2 was conducted within a relatively homogeneous class of systematic musicology students. Although this sample still displayed some variability in the individual amounts of musical training, one extreme of the distribution from Experiment 1 (i.e., professionals musicians) was not represented. Therefore, it is possible that limited variation in musical training in Experiment 2 influenced variability in imagery and SMS, with reduced variability resulting in diminished covariation between imagery and SMS ability. Some support for this comes from the finding that imagery/training relationships that have been commonly reported in the literature, and were also observed in Experiment 1, were found to be non-significant in Experiment 2 (cf. Janata & Paroo, 2006; Aleman et al., 2000). To clarify these apparent discrepancies, we plan to supplement the sample employed in Experiment 2 with both musical experts and non-musicians in the future.

Notwithstanding the above issue, the results of Experiment 2 extend our current knowledge on imagery/synchronization relationships by showing that temporal imagery ability is more closely related to performance in simple SMS tasks than is pitch imagery ability. In addition, our findings suggest that pitch and temporal imagery are two dissociable dimensions that can develop differently in musicians, possibly depending on the importance of each domain for the respective instrument (e.g., piano, where pitched are fixed, versus violin, where pitches are variable). The notion of independent pitch and temporal imagery is consistent with models of music perception that assume that musical input is analyzed by two parallel and largely independent subsystems specialized for dealing with either pitch or temporal information (e.g., Peretz & Coltheart, 2003). Considering that auditory images have been found to share many attributes with their corresponding percepts (e.g., Intons-Peterson, 1992), the functioning of these two subsystems in imagery might be similar to perception, with one subsystem dealing with the melodic content (i.e., producing successive pitches) while the other system focuses on the temporal content (i.e., producing successive durations).

The findings of the current study generally support the idea that auditory imagery abilities are of importance for interpersonal action coordination in musical ensembles. Musicians may establish and maintain synchrony by using their imagery ability to predict the time course of each others' actions. Nevertheless, we cannot completely rule out that other aspects like general musical ability (independent of music training), domain-general working memory capacity and intelligence, or the interaction of domain-specific cognitive and motor predictive processes might also have a contribution to the observed relationship between auditory imagery abilities and SMS accuracy. Future studies will be designed to further clarify this issue.

ACKNOWLEDGMENT

We thank Kerstin Traeger, Juliane Zeiss and Andreas Weber for their help with data acquisition. The research reported in this article was supported by the Max Planck Society.

REFERENCES

Aleman, A., Nieuwenstein, M. R., Boecker, K. B. E., & de Haan, E. F. H. (2000). Music training and mental imagery abilities. N*europsychologia, 38*, 1664-1668.

Deutsch, D., & Pierce. J. R. (1992). The climate of auditory imagery and music. In Reisberg, D. (Ed.), *Auditory Imagery* (pp. 237-260). Hillsdale: Lawrence Erlbaum.

Halpern, A. R., & Zatorre, R. J. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron, 47*, 9-12.

Herholz, S. C., Lappe, C., Knief, A., & Pantev, C. (2008). Neural basis of music imagery and the effect of musical expertise. *European Journal of Neuroscience, 28*, 2352-2360.

Highben, Z., & Palmer, C. (2004). Effects of auditory and motor mental practice in memorized piano performance. *Bulletin of the Council for Research in Music Education, 159*, 58-65.

Intons-Peterson, M. J.(1992). Components of auditory imagery. In Reisberg, D. (Ed.), *Auditory Imagery* (pp. 45-72). Hillsdale: Lawrence Erlbaum.

Janata, P., & Paroo, K. (2006). Acuity of auditory images in pitch and time. *Perception & Psychophysics, 68*, 829-844.

Keller, P. E. (2008). Joint action in music performance. In Morganti, F., Carassa, A., & Riva, G. (Eds.), *Enacting Intersubjectivity: A cognitive and social perspective to the study of interactions* (pp. 205-221). Amsterdam: IOS Press.

Pechmann, T., & Mohr, G. (1992). Interference in memory for tonal pitch: Implications for a working-memory model. *Memory & Cognition, 20*, 314-320.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6*, 688-691.

Repp, B. H. (2002). The embodiment of musical structure: Effects of musical context on sensorimotor synchronization with complex timing patterns. In Prinz, W., & Hommel, B. (Eds.), *Common mechanisms in perception and action: Attention and Performance* (pp. 245- 265). Oxford, U.K.: Oxford University Press.

Schendel, Z. A. & Palmer, C. (2007). Suppression effects on musical and verbal memory. *Memory & Cognition, 35*, 640-650.

Trusheim, W. H. (1991). Audiation and mental imagery: Implications for artistic performance. *The Quarterly Journal of Music Teaching and Learning, 2*, 138-147.