

Self versus other in piano performance: detectability of timing perturbations depends on personal playing style

Bruno H. Repp · Peter E. Keller

Received: 4 July 2009 / Accepted: 24 November 2009 / Published online: 11 December 2009
© Springer-Verlag 2009

Abstract Differences between recorded repetitions of one's own movements are detected more readily than are differences between repetitions of others' movements, suggesting improved visual discrimination due to heightened resonance in the observer's action system and/or relatively accurate internal action simulation (Daprati et al. in *Conscious Cogn* 16:178–188, 2007). In Experiment 1, we attempted to replicate this finding in the auditory modality. Pianists were recorded playing musical excerpts three times and later judged whether pairs of recordings were the same take or different takes of the same excerpt. They were no better at distinguishing different takes of their own playing than those of other pianists' playing, even though discrimination and self-recognition were well above chance. In Experiment 2, the same pianists tried to detect small local timing deviations that had been introduced artificially. They were better at detecting such deviations in their own performances than in those of another pianist, but only if the deviations were placed at points of a pre-existing self-other difference in local timing. In that case, pianists' ability to predict their own characteristic action pattern did aid their perception of temporal irregularity. These results do not support the perceptual sharpening hypothesis of Daprati et al. in the musical domain, but they do suggest that pianists listening to performances generate idiosyncratic temporal expectations, probably through internal action simulation.

Keywords Self-recognition · Action simulation · Motor resonance · Expectancy · Timing perception · Music performance · Forward models

Introduction

Following the seminal discovery of mirror neurons in monkeys (for a review, see Rizzolatti and Craighero 2004), many studies have demonstrated that human action observation also engages some of the same neural systems that are involved in carrying out actions like those observed (e.g., Calvo-Merino et al. 2005; Cross et al. 2006; Haslinger et al. 2005; Lahav et al. 2007). The observer's action system seems to resonate to the observed action (Schütz-Bosbach and Prinz 2007a; Viviani 2002) and predictions about the action's future course are generated on the basis of an internal simulation process (Jeannerod 2003a, b; Schütz-Bosbach and Prinz 2007b; Wilson and Knoblich 2005). The closer an observed action matches the observer's characteristic movement pattern, the stronger the resonance of his or her action system and the more accurate the simulation will be (Knoblich and Flach 2003).

The closest match obviously occurs when an observer watches his or her own action. Visual observation of one's own actions occurs in daily life when carrying out hand movements under visual guidance or when watching one-self move in a mirror. In the auditory domain, where listening to sounds produced by human movement amounts to action observation, self-observation is common, especially in speaking, singing, and instrumental music performance. Observers also often recognize their own movements when they are presented off-line as a video or audio recording, even when they have not seen or heard these specific actions previously. For example, participants presented

B. H. Repp (✉)
Haskins Laboratories, 300 George Street,
New Haven, CT 06511-6624, USA
e-mail: repp@haskins.yale.edu

P. E. Keller
Max Planck Institute for Human Cognitive
and Brain Sciences, Leipzig, Germany

with point-light displays of filmed actions can identify themselves at levels much better than chance, even though they would hardly ever have observed themselves carrying out those actions in a mirror, let alone as light points (Loula et al. 2005; Sevdalis and Keller 2009). Repp and Knoblich (2004) had pianists listen to MIDI recordings made several months previously, half of which were made on a silenced piano; yet the pianists often could identify their own playing, even if they had not heard the performance previously. Repp and Knoblich argued that self-recognition was based on feelings of familiarity derived from the increased resonance of the action system to the pattern of one's own actions and/or on the absence of mismatches between expectations generated by an internal simulation of the performance and the observed performance. The timing pattern of the performances seemed to be particularly important, for self-recognition remained intact after all other cues (except for articulation, also a temporal cue) were eliminated.

Heightened resonance in the action system or relatively accurate expectancies may not only lead to explicit self-recognition but also to enhanced sensitivity to small differences between movement patterns. Daprati et al. (2007) demonstrated this in a study that directly inspired the present research. They showed participants pairs of video clips of a virtual hand (a stick figure) performing simple actions (tapping, finger flexion, line tracing). A virtual hand was used to eliminate any morphological cues to hand identity. The movements of the virtual hand were modeled after video recordings of a participant's own hand or of another participant's hand. The two videos in a pair always showed the same kind of action, but they were either (1) identical (the same clip played twice, performed by self or other) or (2) showed the action carried out by the same person (self or other) on different occasions or (3) showed the action carried out by different persons (self and other). The task was to decide quickly whether the two videos were identical or not. Participants were both faster and more accurate in responding "different" to the second type of trial when it showed their own movement than when it showed someone else's. They were not required to discriminate explicitly between self and other, and reported informally that they rarely felt having seen their own movements. Daprati et al. considered two not mutually exclusive explanations of their findings. First, participants' greater familiarity with their own styles of movement may have enabled them to detect finer differences in execution. Second, action observation may have induced motor resonance (see Iacoboni et al. 2001) and/or engaged an internal forward model (see, e.g., Wolpert and Ghahramani 2000), with the resonance being stronger—and the predictions generated by the internal model more

precise—for one's own actions. The latter would be the case because, in the absence of specific information about and familiarity with another person's action patterns, a forward model (i.e., action simulation) essentially predicts how the observed action would be carried out by the observer's own action system.

We wondered whether we could replicate the findings of Daprati et al. (2007) in the auditory domain, specifically in music performance. In Experiment 1, we used a design similar to that of Daprati et al. to investigate whether musician listeners would be more sensitive to differences among their own performances than to differences among someone else's performances. Experiment 2 took a different approach to address the related question of whether listeners would be more sensitive to perturbations of their own recorded action pattern than to perturbations of someone else's action pattern.

Experiment 1

The design of Experiment 1 followed Daprati et al. (2007) fairly closely, using piano performances instead of hand movements. There were a few methodological differences: first, we did not include pairs of performances by different pianists (self and other) because they would have been too easy to discriminate, given differences in tempo, timing, dynamics, and articulation (e.g., *staccato* vs. *legato*). We think that in the study of Daprati et al., too, the self-other pairs of videos were not really essential. Second, our design was completely counterbalanced in that pairs of pianists listened to identical materials, with "self" for one pianist being "other" to the other pianist, and vice versa. (The design of Daprati et al. was only partially counterbalanced in this way.) Third, we asked participants to identify their own performances as well as judge differences between performances. Thus, our participants were aware that some of the performances were their own, but we did not expect this to affect their discrimination judgments. If anything, we thought it might make them listen more carefully to all performances, so as to determine which were their own.

Methods

Participants

The participants were 12 pianists (9 women and 3 men, ages 22–30) who resided in Leipzig, Germany, but were mostly from East Asia. They included both advanced students and professionals, had played the piano since ages 4–8, agreed to come for multiple sessions, and were paid for their services.

Recordings

In the first session, each pianist was recorded playing twelve musical excerpts three times on a Yamaha Clavinova CLP 150 digital piano. The excerpts included beginnings of preludes by Bach and of sonata movements by Mozart, Beethoven, and Schubert, three by each composer, about 20 s in length on average. (These were the same excerpts as used in Repp and Knoblich 2004; for details, see their Table 1.) The pianists were sent the musical scores of the excerpts in advance and had the opportunity to practice them before coming to the laboratory. During the session, a tempo was suggested for each excerpt by means of a metronome. The participant practiced the excerpt with the metronome running until he or she felt comfortable with the music and then played it with the metronome turned off three times in succession. The pianists were instructed to play the music the same way each time. The performances were recorded as MIDI files. At the end of the recording session, the pianists were asked to indicate for each excerpt whether they had played it previously, knew it from listening, or were unfamiliar with it (prior to receiving the musical scores for the study). If these responses are scored as 2, 1, and 0, respectively, mean familiarity scores ranged from 0.42 to 1.58 across excerpts, and from 0.33 to 1.75 across pianists.

Because the recorded performances were not entirely free of errors, two musically trained research assistants and author BHR listened to all performances, marked any errors they could hear, and then made the necessary corrections in the MIDI files. Errors included wrong notes, extra notes, and obviously unintended anomalies of timing and articulation that were specific to individual takes and thus would have made discrimination among takes too easy. Well over 100 such errors were corrected.

Materials

Six pairs of pianists were formed in an arbitrary fashion. For each pair, 48 trials were created from these two pianists' performances of the 12 excerpts, four trials for each excerpt. Each trial contained two performances of an excerpt played by the same pianist. These two performances were either identical (i.e., one of the three takes, chosen arbitrarily, was repeated) or different (the other two takes), and they were played by one or the other pianist.

Procedure

Pianists were tested individually several months after the recording session. A program written in MAX/MSP controlled the procedure. The MIDI files were played back on the same digital piano on which they had been recorded,

and pianists listened over Sennheiser HD270 earphones. The trials were different and in a different random order for each pair of pianists, but the two pianists in a pair received exactly the same random order of the same trials. They were asked to judge after each trial whether the two performances had been identical or whether there had been any difference whatsoever, and then they reported whether they thought the performances were by themselves or by another pianist. They gave these responses by clicking virtual buttons on a computer screen using the mouse. The second response triggered the next trial after a delay of 2 s. The two performances in a trial were separated by 2 s of silence. The pianists were informed that "same" and "different" trials were equally frequent, as were "self" and "other" trials. They were encouraged to give equal numbers of responses in each category and to guess when not sure.

Results

The discrimination task was far from easy, but participants clearly performed above chance level (50%), with 64.2% correct responses overall, $t(11) = 7.43$, $P < 0.001$. They were also able to identify their own performances at better than chance level, with 67.0% correct responses overall, $t(11) = 5.09$, $P < 0.001$. However, contrary to our hypothesis, they were not significantly better at discriminating identical from non-identical performances when the performances were their own rather than another pianist's, 62.5 versus 66.0% correct, $t(11) = -0.73$, n.s. (Clearly, the conclusion would be the same if we computed d' instead of percent correct.)

Closer inspection of the data revealed a very high false alarm rate: the pianists responded "different" to performances that were in fact the same 40.6% of the time. However, it could be argued that our hypothesis that discrimination would be sharpened by heightened resonance in the listener's action system and/or relatively accurate simulation does not really apply to identical pairs: There is no good reason why differences that do not exist should be "perceived" better in one's own performances. (Alternatively, or in addition, false alarms may not reflect perception but simply a bias to respond "different," which had been encouraged by the instructions to give equal numbers of "same" and "different" responses.) Therefore, we examined correct responses to pairs of non-identical performances (i.e., hits) separately. However, the results were similar, 67.4 versus 70.8% for self versus other. We also repeated this analysis with pairs of pianists rather than individual pianists as the units because if one pianist's performances were much easier to discriminate than the other pianist's in a pair, this would have inflated the variability of the self-other difference at the individual level.

However, the results were no different: Of the six pairs of pianists, only three exhibited a difference in the predicted direction.

Although the findings of Daprati et al. (2007) suggest that explicit self-recognition is not a prerequisite for a self-other difference in perception, it is not unreasonable to suppose that a listener's action system is most strongly engaged in those trials in which self-recognition occurs (correctly or not). Therefore, we also looked separately at trials that received "self" responses. However, they again did not show the predicted difference between true self and true other, 59.5 versus 63.0% correct. Paradoxically, only trials judged as "other" showed a slight difference in the predicted direction, 67.5 versus 64.1% correct for true self versus true other (average of individual percentages, with one participant excluded because of an empty cell), but it was far from significance, $t(10) = 0.95$, n.s.

Discussion

Pianists were no better at discriminating different takes of their own performances than of others' performances. Thus, we failed to replicate the results of Daprati et al. (2007) in the auditory domain. The reasons for this are not entirely clear. It seems unlikely that any of the differences in design was responsible. The absence of pairs of performances by different pianists, which would have been easy to discriminate, could hardly have made a difference with regard to self versus other; if they had been included, the percentage of correct discrimination of non-identical performances by the same pianist would probably have been much closer to chance. The complete counterbalancing, achieved by pairing all pianists, can only have been advantageous compared to the study of Daprati et al. Finally, there is no good reason why the explicit identification of performances as self or other should have had any influence on the discrimination judgment that preceded it. Daprati et al. took pains to prevent even implicit self-recognition (if their participants asked before the experiment whether their own movements had been included among the stimuli, they gave them a negative answer), but it is not clear why this was necessary.

More likely, the crucial difference between the studies lies in the nature of the materials. First, our stimuli were much longer than theirs (about 20 vs. 2.7 s) and thus made much greater demands on memory. Indeed, it is quite unlikely that pianists were able to remember all details of the first performance in a pair, and their strategy was probably to listen for any unusual local features of the first performance and keep those in mind when listening to the second performance. Second, whereas the stimuli of Daprati et al. (2007) could be discriminated only on the basis of a space-time trajectory, our stimuli could be discriminated

along a number of dimensions, including tempo, timing, dynamics, and articulation. These dimensions may not be equally relevant to the hypothesis being tested. For example, global differences in tempo or dynamics (loudness), which are not very agent-specific, seem less relevant to the self-other distinction than the detailed patterns of timing, dynamics, and articulation, which reflect an artist's personal style more strongly.

It is possible that timing is really the crucial dimension that elicits motor resonance and that is simulated by internal forward models of music performance. Flach et al. (2004) found that individual timing provides sufficient information for self-recognition of rhythmic hand clapping, and Repp and Knoblich (2004) showed that differences in expressive timing and articulation are sufficient for self-recognition in piano performance, with timing probably being more important than articulation. If the present performances were discriminated mainly on the basis of global differences in tempo, dynamics, or articulation, the participants' responses would not have reflected the different resonance strengths or local expectancies that the timing patterns elicited in their action systems. This idea could be tested by presenting performances that differ only in timing, a manipulation comparable in its stringency to the use of a virtual hand by Daprati et al. (2007). It is doubtful, however, whether participants would perform above chance level with such stimuli, given the length of the excerpts, the great similarity in timing of repeated performances of the same excerpt by the same pianist with the same interpretative intentions (Palmer 1989; Repp 1995), and the difficulty of remembering temporal detail. Therefore, we decided to take a different approach in Experiment 2, introducing timing differences artificially and changing the task from discrimination to detection.

Experiment 2

In this experiment, the task was to detect artificially introduced deviations from the original expressive timing pattern (timing profile) of single performances. The hypothesis was that such deviations would be easier to detect in one's own performances than in someone else's. The theoretical argument remains the same: one's own performance should engage the action system more than someone else's performance does, and in particular the temporal predictions of an internal forward model should be more accurate. The deviations to be detected should be perceived as violations of these predictions. In one's own performance, such violations would be due mainly to the artificially induced deviations, whereas in listening to someone else's performance there might be other violations of expectancies due to individual differences in

preferred expressive timing patterns, which would serve as distracters and generate false alarms.

To distinguish between a generally heightened sensitivity to deviations in one's own performance (the perceptual sharpening hypothesis derived from Daprati et al. 2007) and a sensitivity that is more specifically tied to individual performance styles, we introduced two types of deviation, as explained in more detail in the “Methods”. We again formed pairs of pianists and examined their performance timing patterns beforehand. Type 1 (“neutral”) deviations were made to occur at points at which the two pianists' performances did not differ, whereas Type 2 (“biased”) deviations made the two pianists' performances more similar to each other at a point in time where they originally differed. Type 1 deviations thus tested the hypothesis that any timing deviation would be easier to detect in one's own performances than in someone else's, whereas Type 2 deviations tested the more specific hypothesis that deviations in one's own performance (which necessarily violate expectations) would be easier to detect than deviations in another pianist's performance that happen to meet one's own local expectations. Although this second hypothesis may seem trivial, it should be remembered that the two pianists in a pair listened to the same stimuli, with only the roles of self and other being interchanged. Therefore, a greater self-other difference in detection of Type 2 deviations than in detection of Type 1 deviations would constitute specific evidence that internally generated predictions reflect individual performance styles.

The experiment had two parts. Both parts involved the same stimuli, but whereas in Part 1 only detection responses were collected, in Part 2 we also measured reaction time (RT) and asked participants to identify their own performances. Memory for specific locations of perceived deviations in the musical excerpts was highly unlikely, so that Parts 1 and 2 can be considered independent.

Methods

Participants

The participants were the same as in Experiment 1, and they were paired in the same way.

Materials

We selected one performance (the last of the three takes) of each excerpt recorded by each pianist. In each performance we introduced four local timing deviations, two of Type 1 and two of Type 2. All timing deviations were increments of inter-onset intervals (IOIs) between successive notes because previous research (Repp 1998a, b) had suggested that detection of IOI increments (perceived as hesitations)

tends to reflect listeners' temporal expectations more strongly than detection of IOI decrements (perceived as hastenings). This may be so because in expressive performance it is more common to linger than to hasten.

To decide where to insert the deviations, we first derived the *timing profiles* of all selected performances from the MIDI data. The timing profile is a series of IOIs corresponding to a constant note value. Loosely speaking, these IOIs represent the performed durations of successive eighth notes or sixteenth notes, depending on the excerpt. Whenever the onsets of several notes coincided in the score, we took the onset of the note in the most important voice as the marker in the performance. Whenever an IOI corresponded to a multiple of the base note value (e.g., a quarter-note), we divided it by the number of base note values it contained. Then we normalized all IOIs by dividing them by the mean IOI. Once we had the normalized timing profiles of the two performances of the same excerpt by the two pianists in a pair, we computed their difference, resulting in a *difference profile*. The difference profile showed values near zero where the two pianists' local timing differed little, and negative or positive differences where the two pianists differed in their local timing. These local timing differences typically occurred in a single IOI and thus showed up as “spikes” in the difference profile.

We placed the Type 1 deviations at points where the difference profile was close to zero (at two different points in each pianist's performance), the Type 2 deviations for one pianist (A) at two points where the difference profile showed negative spikes (showing that pianist A played faster at this point than pianist B, if the difference profile represents $A - B$), and the Type 2 deviations for the other pianist (B) at two points where the difference profile showed positive spikes (showing that pianist B played faster at this point than pianist A). In selecting appropriate points for introducing deviations, we kept a minimum distance of one measure between successive deviations in the same performance and also tried to avoid placing deviations at corresponding points within different measures. Figure 1 gives one illustrative example, explained in the caption.

The magnitude of all deviations introduced was 20% of the IOI in the original performance.¹ In most cases, the changed IOI represented the base note value, but occasionally it was longer; only very rarely did it represent a rest. Each deviation was implemented by proportionally stretching the intervals between all successive MIDI events

¹ A change of 20% seems large, but local deviations from expressive timing profiles are not easy to detect. From Repp's (1998b) data a mean detection (hit) rate of about 70% can be predicted for hesitations of that magnitude, but in that study a single musical excerpt was presented repeatedly. With many different excerpts, each presented only once, the detection rate was expected to be lower.

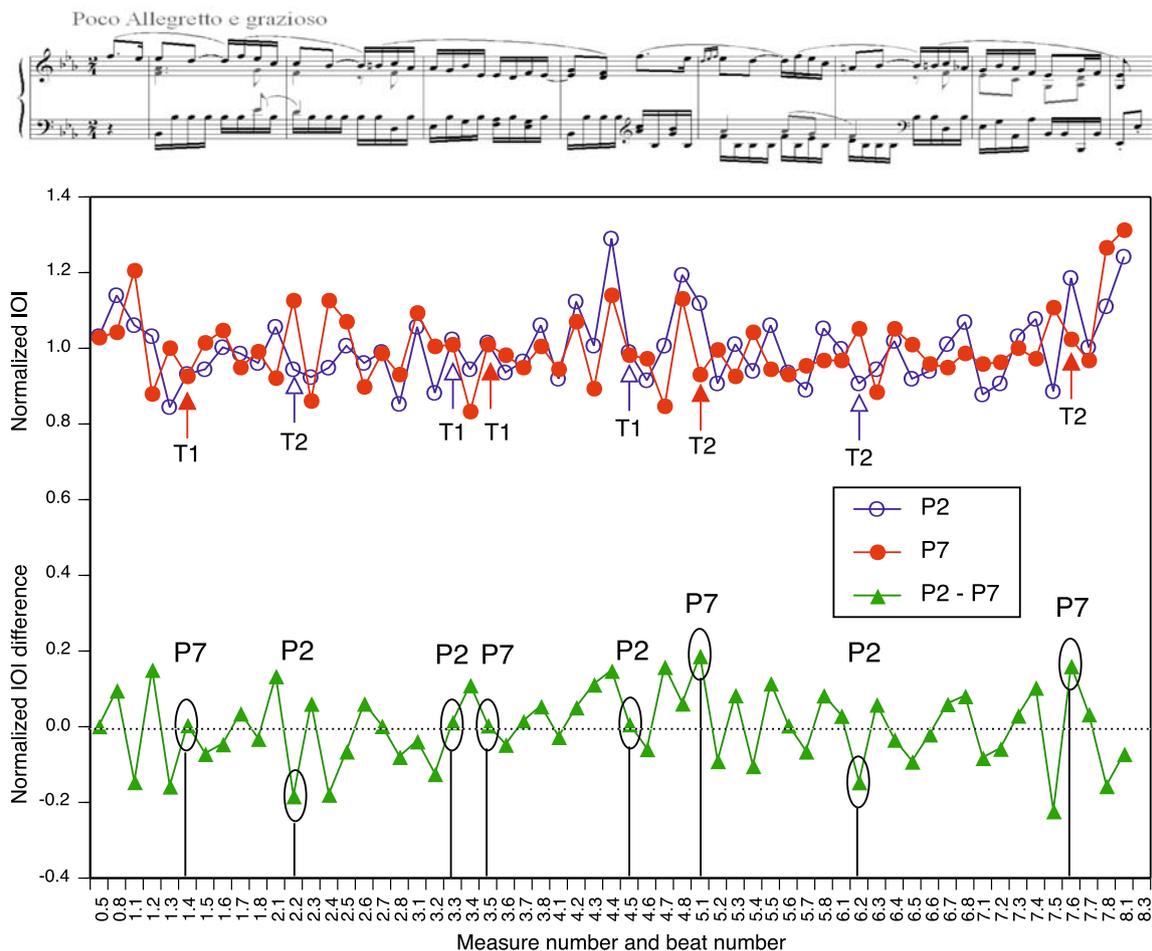


Fig. 1 Illustration of how local deviations were created in one musical excerpt (the beginning of the fourth movement of Beethoven's Sonata in E-flat major, op. 7) for two pianists in a pair (P2, P7). The two functions on *top* are the normalized inter-onset intervals (IOIs) or timing profiles of the two pianists, whereas the function at the *bottom* is the difference timing profile (P2 – P7). *Ovals* indicate points in the difference profile that were selected for insertion of deviations. *Vertical lines* point to the measure and beat numbers of the events terminating the IOIs to be changed. *Labels above the ovals* refer to the pianist in whose timing profile the change was made.

(note onsets and offsets; there was no pedaling) occurring within the to-be-changed IOI. The event onset times of the modified performances were reconstituted by cumulating the intervals between successive MIDI events. All this was done in a spreadsheet program by author BHR who also listened to all modified performances to make sure there were no editing mistakes.

Procedure

The pianists returned for Experiment 2 several months after Experiment 1. The session had two parts. In Part 1, the 24 modified performances (12 excerpts played by self or other) were played in a random order. Because the task was

expected to be difficult, each performance was immediately repeated once to give participants two opportunities to detect hesitations. Participants had a 24-page response booklet showing the musical scores of the excerpts, a separate page for each trial. They were told that there were up to four hesitations in each performance and were asked to circle the prolonged note(s) in the score. The instructions explained that the hesitations had been introduced artificially and sounded inappropriate, as if the piano keys had gotten stuck for a brief moment. Participants were not told that there were always exactly four hesitations, in order to discourage random guessing.

In Part 2, the same 24 performances were again presented twice, but now in two separate blocks (without repeats

within blocks). The random order was the same in the two blocks. Participants were instructed to press the “down arrow” key on the computer keyboard as quickly as possible whenever they heard a hesitation. In addition, after listening to each excerpt, they were asked to indicate whether the performance was by her/himself or by another pianist by clicking one of two buttons on the computer screen. A customized MAX program registered RTs (measured from the end of each changed IOI) and identification responses.

In both parts of the experiment, the two pianists in a pair listened to exactly the same random sequences of performances, which were reproduced on the digital piano and heard over earphones.

Results

Part 1

We scored responses as hits if participants circled the correct note or an adjacent note. (A tendency to circle the following note, in particular, is common in such detection tasks; see Repp 1998a, b.) False alarm rates were estimated as the mean number of incorrect responses per excerpt divided by 4 and multiplied by 100 to obtain a percentage.² Two participants stood out in that they had unusually low hit rates together with very high false alarm rates. We excluded these participants' data because we suspected they had misunderstood the instructions.³

Overall detection performance was 41.4% correct (hits). A two-way repeated measures ANOVA revealed a significant main effect of deviation type, $F(1,9) = 34.46$, $P < 0.001$: Type 1 deviations (50.2% correct), which violated both pianists' timing expectations, were easier to detect than Type 2 deviations (32.5% correct), which were designed to violate only one pianist's expectations. The main effect of self versus other was not significant, but the interaction was reliable, $F(1,9) = 6.36$, $P = 0.033$: Type 2 deviations were detected more easily in performances by self than by other (37.5 vs. 27.5% correct), whereas Type 1 deviations showed a small difference in the opposite direction (48.3 vs. 52.1% correct).

² First, we determined the mean number of positions in the music where a deviation could occur. Across the 12 excerpts, with the initial and final two positions excluded, this number was 63. Given that four deviations occurred in each excerpt and a lenient scoring criterion (± 1 position) was used, there was an average probability of $12/63 = 0.19$ that a random response would be scored as a hit. In other words, about 1/5 of the music consisted of “signal zones” and 4/5 consisted of “noise zones.”

³ Like most other participants, they were not native speakers of German and presumably failed to understand that they were to listen for artificially introduced hesitations. They probably responded instead to local expressive slowing, which of course can be detected if it is large enough and if such detection is the listener's intention.

The mean false alarm rates for performances by self and other were 11.7 and 12.7%, respectively. These rates are much smaller than the hit rates, which goes to show that detection of deviations was clearly better than chance (except for the two excluded participants). Given that false alarms cannot be distinguished according to deviation type and that the rates were similar for self and other, there was no point in calculating d' indices.

Part 2

We accepted responses with RTs between 200 and 1,500 ms as hits. The two participants who had been excluded from analysis in Part 1 were excluded again. (They again stood out through their high false alarm rates.) The overall percentage of hits was 35.4%. The ANOVA on hit percentages replicated the findings of Part 1: Type 1 deviations were detected more often (45.6%) than Type 2 deviations (25.1%), $F(1,9) = 33.07$, $P < 0.001$, and the interaction between self-other and deviation type was also significant, $F(1,9) = 10.53$, $P = 0.010$. Again, Type 2 deviations were detected more readily in performances by self than by other (29.2 vs. 21.0% correct), whereas Type 1 deviations showed a small difference in the opposite direction (44.2 vs. 47.1% correct).

Analysis of RTs revealed only a nearly significant main effect of deviation type, $F(1,9) = 4.73$, $P = 0.058$: mean RTs were shorter for Type 1 than for Type 2 deviations (557 vs. 614 ms). The pattern of the interaction, though not statistically reliable, $F(1,9) = 2.14$, $P = 0.177$, paralleled the one for hits: mean RTs for Type 2 deviations were shorter for self than for other (589 vs. 639 ms), whereas mean RTs for Type 1 deviations showed an opposite difference (575 vs. 539 ms).

False alarm rates were quite low in Part 2 (apart from the two excluded participants), 4.0% for self and 3.1% for other, and even these are overestimations because the number of false alarm opportunities was actually greater than in Part 1 (but was assumed to be four per performance in the calculation). It seems that participants adopted a more conservative response criterion when they had to respond immediately.

As in Experiment 1, participants were able to identify their own performances at better than chance levels, 63.7% correct, $t(11) = 4.29$, $P < 0.001$, though no better than in Experiment 1. There was also no improvement in self-identification between Blocks 1 and 2 (64.3 vs. 63.2% correct).

Discussion

In this experiment, we were able to demonstrate a perceptual advantage for self over other. We attribute this

success in part to the fact that perception was restricted to temporal differences. However, the advantage occurred with only one of the two types of temporal deviation.

Let us consider first the result that Type 1 deviations were generally easier to detect than Type 2 deviations, which is not difficult to explain. Type 1 deviations occurred at points where two paired pianists' performances did not differ with regard to timing. This means that, in most cases, a momentary slowing probably would not be expressively appropriate at these points. Therefore, most of the Type 1 deviations were likely to be perceived as true hesitations in the music by both pianists. In contrast, Type 2 deviations usually occurred at points where one pianist showed expressive slowing but the other pianist did not (see Fig. 1). Assuming the slowing was intended, this suggests that expressive slowing is aesthetically acceptable at that point.⁴ Therefore, a Type 2 deviation in a performance that did not originally show slowing at that point might not be perceived as a disruption of the musical flow but rather as a valid expressive gesture and therefore would be difficult to detect. (A Type 2 deviation was never introduced at a point at which slowing was already present.) Occasionally, Type 2 deviations were introduced at points where one pianist showed a local hastening but the other pianist did not. (Points were chosen on the basis of spikes in the difference profile, not of the timing of the individual performances at that point.) Local hastening rarely occurs for expressive reasons and thus may have been unintended, in which case the Type 2 deviation merely corrected a timing error and restored local temporal regularity. Then it was difficult to detect for that reason.

Type 1 deviations were not easier to detect in one's own performance than in another pianist's performance, although such a difference was predicted by the hypothesis that perception is enhanced by stronger motor resonance (Daprati et al. 2007). Thus, Experiment 2, like Experiment 1, fails to support this hypothesis, which apparently does not apply straightforwardly to music perception. The negative result also indicates that pianists were not simply more alert or more critical when listening to their own performances (which they quite often did not recognize).

Only Type 2 deviations were easier to perceive in one's own performance than in another pianist's. When heard in the context of one's own performance, a Type 2 deviation is not so much a hesitation as an expressive slowing that is

atypical of one's own style of playing. Thus, it creates a mismatch with an internal prediction of the expressive timing and is—we presume—detected on the basis of that mismatch, not (or only rarely) because it disrupts the musical flow. For the other pianist, who did not generate the performance but showed expressive lengthening in her own performance at this point, the Type 2 deviation should be quite difficult to detect because it meets her local expectations. Detection could occur only on the basis of an incongruity or discontinuity with local context, or because the precise magnitude of the deviation does not match expectations. Such local cues must have played a role because the difference between self and other in detection of Type 2 deviations was not very large. The absence of a difference in false alarm rates between self and other and the generally low false alarm rates (especially in Part 2 of the experiment) also reveal that expressive slowing, when it occurred in its proper undisturbed context, was rarely perceived as a hesitation, even when it did not match the listener's expectations. Another factor contributing to the small size of the self-other difference in detection of Type 2 deviations is that some of these deviations may have merely corrected a timing error (local hastening) and then presumably were equally difficult to detect by self and other.

Even though participants heard the same performances four times in Experiment 2, their self-identification did not improve and remained at a modest level of accuracy. This is probably because the chosen excerpts encouraged relatively "straight" playing with only small expressive timing deviations. Thus the cues for self-identification were relatively limited, even though they included tempo, dynamics, and articulation. Timing, however, is likely to be the most important dimension for self-recognition (Flach et al. 2004; Repp and Knoblich 2004).

General discussion

In this study, we attempted to replicate and extend the findings of Daprati et al. (2007) using musical materials. Experiment 1 was similar in design to theirs, but our auditory materials were much more complex and longer in duration than the simple visual movement trajectories they displayed. We failed to find a self-other difference in discrimination of different takes of the same excerpt, perhaps because takes differed not only in timing but also along other dimensions that may be less relevant to the self-other distinction. Resonance in the listener's action system, being a dynamic process that unfolds in time, may be most closely linked to timing, especially if it is not accompanied by vivid auditory imagery. However, timing differences between takes were probably too small to be detectable.

⁴ To determine whether local features of the selected performance of each excerpt were intended or not, we could have analyzed the timing profiles of all three renditions of each excerpt by each pianist and examined their consistency. However, this would have added much time to what was already an extremely time-consuming process of stimulus preparation. Thus, some proportion of Type 2 changes may have been introduced at points of unintended deviation in local timing.

Tempo, dynamics, and articulation probably provided more salient cues to differences between takes, and given the high demands that our materials placed on memory, participants probably compared just the beginnings and ends of takes, and also listened for any remaining odd features of performances that we had been unable to correct during stimulus preparation.

Experiment 2 focused on timing exclusively and used a detection task similar to that in Repp (1998a, b). The results showed that artificially introduced timing deviations are detected more readily in one's own performance than in someone else's, but only if they occur at points where the two paired pianists' performances differ. If they occur at points where similar temporal expectations are entertained by self and other, there is no self-advantage, contrary to the perceptual sharpening hypothesis of Daprati et al. (2007), which predicts better detection of deviations in one's own performance than in someone else's performance if these deviations are equally unexpected. Although timing deviations must be detected in all cases on the basis of whether they do or do not match internally generated temporal expectations, it is only at points where the two pianists have different expectations that a self-other difference emerges. In that case, the deviation in one's own performance matches one's own expectation less than it matches the other pianist's expectation, and therefore it is easier to detect in one's own performance. Thus, we find evidence of performer-specific temporal prediction (i.e., a perceptual bias, equivalent to a direction-specific increase in sensitivity), but no sharpening of overall perceptual sensitivity when listening to one's own performance.

One reason why we failed to find perceptual sharpening for one's own performance may be the different ecological status that music performances have compared to visually observed actions. Daprati et al. (2007) argue that fine details of one's own actions need to be perceived in order to correct errors and improve performance (which is undoubtedly also true for music), but that the actions of others are perceived mainly in terms of their goals and intentions, for which coarser perceptual resolution will suffice. In the case of music, however, the performances of others rarely convey goals or intentions that go beyond the music itself: The perfect production of the music is the goal. Therefore, musicians listen to performances by others just as critically as they listen to their own performances. Another possibility is that the additional task of distinguishing self from other somehow interfered with the detection of subtle performance differences, although we find that unlikely.

The self-advantage in sensitivity to Type 2 timing deviations found in the current study complements previous work on the self-other distinction in music performance. Collectively, this research highlights the importance of temporal information in musical action

simulation. To date, three studies (including the current one) have examined self-recognition with musical recordings, and all have yielded results that are consistent with the notion that expressive timing provides clues to the identity of the agent who performed the action. In the earliest study, Repp and Knoblich (2004) demonstrated the sufficiency of timing cues for the self-recognition of solo piano performances by systematically removing other cues from the recordings. Taking a different approach, Keller et al. (2007) investigated the role of action simulation in temporal prediction by examining the relationship between self-recognition and synchronization in piano duets. Keller et al. found that pianists were not only able to distinguish between their own and others' recordings of one part of a duet, but they were also better able to play the other part in synchrony with their own recordings. Moreover, self-recognition was positively correlated with the self-advantage in synchronization, which suggests that both tasks may be mediated by temporal predictions based on an online simulation process. The findings of the current study suggest that such predictions can also influence perceptual sensitivity, but only when the perceptual targets are designed to exploit differences in individual performance style that evidently are preserved in each individual's temporal expectations. Thus, the quality of the match between observed timing, on one hand, and temporal expectations generated via internal simulation, on the other hand, can affect a broad range of behaviors, including self-recognition, sensorimotor synchronization, and the perception of timing deviations.

In conclusion, the present results are ultimately consistent with the hypothesis of internal simulation during action observation, even if they do not provide further support for the perceptual sharpening hypothesis of Daprati et al. (2007). They extend the still sparse auditory perception data in support of the more general internal simulation hypothesis. Together with the results of our previous work, the current study illustrates that music performance is a fruitful domain in which to investigate action simulation and its role in temporal prediction. Given that music is characterized by a high degree of individuality in performance style (see, e.g., Repp 1992), a particularly inviting area for future research concerns the process whereby one individual learns to simulate another's style. This process, which presumably facilitates the understanding of a performer's expressive intentions and coordination in musical ensembles, may involve training a forward model that is calibrated to the timing of another individual's action system (see Wolpert et al. 2003). Paradigms involving musical imitation and complementary action would be well suited to the investigation of this learning process, which requires much more prolonged exposure to another's performance style than our present participants were given.

Acknowledgments This research was supported by the Max Planck Society and also by National Science Foundation grant BCS-0642506 to BHR. We are grateful to Kerstin Träger for recruiting and scheduling the participants, to Angela Schimming and Regine Steinke for running the experiments and analyzing the data, and to Emily Cross and Günther Knoblich for helpful comments on a draft of the manuscript. We also thank Paolo Viviani and an anonymous reviewer for their suggestions on improving the original manuscript.

References

- Calvo-Merino B, Glaser DE, Grèzes J, Passingham RE, Haggard P (2005) Action observation and acquired motor skills: an fMRI study with expert dancers. *Cereb Cortex* 15:1243–1249
- Cross ES, de Hamilton AFC, Grafton ST (2006) Building a motor simulation de novo: observation of dance by dancers. *Neuroimage* 31:1257–1267
- Daprati E, Wriessnegger S, Lacquaniti F (2007) Knowledge of one's own kinematics improves perceptual discrimination. *Conscious Cogn* 16:178–188
- Flach R, Knoblich G, Prinz W (2004) Recognizing one's own clapping: the role of temporal cues. *Psychol Res* 69:147–156
- Haslinger B, Erhard P, Altenmüller E, Schroeder U, Boecker H, Ceballos-Baumann AO (2005) Transmodal sensorimotor networks during action observation in professional pianists. *J Cogn Neurosci* 17:282–293
- Iacoboni M, Koski LM, Brass M, Bekkering H, Woods RP, Dubeau M-C, Mazziotta JC, Rizzolatti G (2001) Reafferent copies of imitated actions in the right superior temporal cortex. *Proc Nat Acad Sci* 98:13995–13999
- Jeannerod M (2003a) Simulation of action as a unifying concept for motor cognition. In: Johnson-Frey SH (ed) *Taking action: cognitive neuroscience perspectives on intentional acts*. MIT Press, Cambridge, pp 139–163
- Jeannerod M (2003b) The mechanism of self-recognition in humans. *Behav Brain Res* 142:1–15
- Keller PE, Knoblich G, Repp BH (2007) Pianists duet better when they play with themselves: on the possible role of action simulation in synchronization. *Conscious Cogn* 16:102–111
- Knoblich G, Flach R (2003) Action identity: evidence from self-recognition, prediction, and coordination. *Conscious Cogn* 12:620–632
- Lahav A, Saltzman E, Schlaug G (2007) Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *J Neurosci* 27:308–314
- Loula F, Prasad S, Harber K, Shiffrar M (2005) Recognizing people from their movement. *J Exp Psychol Hum Percept Perform* 31:210–220
- Palmer C (1989) Mapping musical thought to musical performance. *J Exp Psychol Hum Percept Perform* 15:331–346
- Repp BH (1992) Diversity and commonality in music performance: an analysis of timing microstructure in Schumann's "Träumerei". *J Acoust Soc Am* 92:2546–2568
- Repp BH (1995) Expressive timing in Schumann's "Träumerei": an analysis of performances by graduate student pianists. *J Acoust Soc Am* 98:2413–2427
- Repp BH (1998a) Variations on a theme by Chopin: relations between perception and production of deviations from isochrony in music. *J Exp Psychol Hum Percept Perform* 24:791–811
- Repp BH (1998b) The detectability of local deviations from a typical expressive timing pattern. *Music Percept* 15:265–290
- Repp BH, Knoblich G (2004) Perceiving action identity: how pianists recognize their own performances. *Psychol Sci* 15:604–609
- Rizzolatti G, Craighero L (2004) The mirror-neuron system. *Annu Rev Neurosci* 27:169–192
- Schütz-Bosbach S, Prinz W (2007a) Perceptual resonance: action-induced modulation of perception. *Trends Cogn Sci* 11:349–355
- Schütz-Bosbach S, Prinz W (2007b) Prospective coding in event representation. *Cogn Process* 8:93–102
- Sevdalis V, Keller PE (2009) Self-recognition in the perception of actions performed in synchrony with music. *Ann N Y Acad Sci* 1169:499–502
- Viviani P (2002) Motor competence in the perception of dynamic events: a tutorial. In: Prinz W, Hommel B (eds) *Attention and performance XIX: common mechanisms in perception and action*. Oxford University Press, Oxford, pp 406–442
- Wilson M, Knoblich G (2005) The case for motor involvement in perceiving conspecifics. *Psychol Bull* 131:460–473
- Wolpert DM, Ghahramani Z (2000) Computational principles of movement neuroscience. *Nat Neurosci* 3:1212–1217
- Wolpert DM, Doya K, Kawato M (2003) A unifying computational framework for motor control and social interaction. *Philos Trans R Soc B* 358:593–602