ORIGINAL ARTICLE

Rüdiger Flach · Günther Knoblich · Wolfgang Prinz

Recognizing one's own clapping: The role of temporal cues

Abstract Do individuals recognize their own clapping? We investigated this question with an off-line paradigm in which the perception of the claps is temporally separated from their generation. In a first experiment, we demonstrated that acoustic recordings of clapping provide sufficient information for discriminating between one's own performance and that of another participant. Self-recognition was still preserved when the claps were reproduced by uniform tones that only retained their temporal pattern. In a subsequent experiment, we verified that the general tempo provides an important cue for self-recognition. Finally, we showed that self-recognition does not rely on concurrent synchronization of actual movements with the auditory sequences. The results are discussed in the light of current theories about perception-action links.

Introduction

A person is listening to a recording of someone playing the piano. She knows that it could be herself practicing a particular piece. But it could also be someone else practicing the same piece. Would she recognize who is playing? The answer seems to depend on the differences between the two performances as well as on the listener's knowledge about her own performance. If the same person hears herself while playing piano, additional factors contribute to self-recognition: The momentary intention to play that piece of music, proprioceptive feedback from the movements, and on-line comparisons

R. Flach · G. Knoblich · W. Prinz
Max Planck Institute for Psychological Research,
Amalienstrasse 33, 80799 Munich, Germany

R. Flach (✉)
Department of Psychology, University College London,
26 Bedford Way, London, WC1H 0AP, UK
E-mail: r.flach@ucl.ac.uk
Tel.: +44-20-76797570
Fax: +44-20-74364276

of intended and actual action effects. In the following, we pursue the hypothesis that the off-line self-recognition of one's own actions is mediated by a subset of the processes that are also used in on-line self-recognition. This hypothesis is motivated by the common coding approach to perception and action planning (cf. Prinz, 1997; Hommel, Müsseler, Aschersleben, & Prinz, 2001).

The core assumption of this approach is that actions are initiated by the activation of a representation of the intended action effect. Moreover, it is assumed that the perception of an action effect automatically activates to some degree the representation of the action underlying its generation. The amount of activation is assumed to be a function of the similarity between the perceived action effect and the representation of the intended action effect. Because the same action system is involved in the generation of an action and its subsequent perception, the activation of an action representation should be maximal when an individual perceives his or her own actions (cf. also Greenwald, 1970). For this reason, a sufficiently high activation of an action representation during action observation could confer on the observed events the quality of being self-generated.

In order to investigate this hypothesis, several studies have already been carried out. They all relied on an off-line paradigm in which the generation of actions and their perception are temporally separated. For this reason, any significant difference between the perception of self-generated and other actions, subsequently termed authorship effect, cannot be attributed to the concurrent generation of the action and its proprioceptive cues. First of all, Knoblich and Prinz (2001) showed that the kinematics of a drawing movement provide sufficient information for recognizing the movement as self-generated. They presented a single moving dot that reproduced the participant's or another participant's drawings of familiar or unfamiliar symbols. Irrespective of the scaling of the overall size or duration of the trajectory, participants could recognize their own actions. When the drawings were reproduced with a constant velocity, however, self-recognition dropped to chance

levels, suggesting that temporal factors are important for distinguishing one's own actions from the actions of another individual.

Knoblich, Seigerschmidt, Flach, and Prinz (2002) subsequently showed that self-generated actions are perceived differently than other actions, even when the task does not explicitly demand self-recognition. In particular, they presented the kinematic reproductions of the first stroke of the figure "2", recorded in a previous session. The strokes had been drawn as part of a complete figure "2" or they had been drawn in isolation. Having generated half of the strokes themselves, participants indicated whether or not the perceived stroke was part of the drawing of a complete figure "2". Judgmental accuracy was above chance when the reproduced movements had been generated by the observing participants themselves, but not when they had been generated by another participant. Post-hoc analyses suggested that inter-individual differences in the maximum velocity mediated this effect.

Finally, Flach, Knoblich, and Prinz (2003) used a synchronization paradigm in order to test whether corresponding effects occur in on-line movement control. Their participants were required to synchronize a button press with the occurrence of local peaks in continuous sinusoidal or zigzag drawing movements, which had been generated by the participants themselves or by another participant in a previous session. If the movements alternated in height, participants' anticipation performance improved with practice, and this improvement was greater for self-generated actions.

In summary, these studies provide converging evidence in support of the assumption that the activation of action representations mediates action perception. Moreover, these studies suggest that sparse kinematic information suffices for off-line authorship effects to occur. This paucity of perceptual information makes it unlikely that these effects simply reflect effects of familiarity or episodic memory.

The first aim of the present study was to extend this evidence by demonstrating that off-line authorship effects can also occur in the auditory modality. The second aim was to demonstrate that such effects can rely on purely temporal information. This prediction is also consistent with the findings of Keele, Pokorny, Corcos, and Ivry (1985) that led them to conclude that temporal interval perception and interval generation share a common modular device. They instructed participants either to tap regularly with the finger or the foot, or to compare the durations of two intervals. There was a significant correlation (.53) across participants between the variability of the tapping performance and the variability of the perceptual judgments. In addition, Ivry and Keele (1989) reported that patients with cerebellar lesions show specific deficits in both tapping and time perception, and Ivry and Hazeltine (1995) found similar Weber fractions for matched time perception and time production tasks. In a similar vein, Treisman, Faulkner, and Naish (1992) found that time perception judgments

and motor responses are similarly affected by particular frequencies of auditory clicks, suggesting that both rely on the same kind of timing device (cf. also Treisman, Faulkner, Naish, and Brogan, 1990). These observations are relevant to the present study, because they suggest that motor processes can contribute to the perception of self-generated action effects.

Experiment 1

A study by Repp (1987) functioned as the starting point for the present experiments. It mainly addressed the question of what kind of information the acoustic properties of clapping reveal about its generation. Repp instructed participants to clap for 10 s at a self-selected rate. At a later time, he presented recordings of ten consecutive claps of each participant in random order and asked the participants, who were all known to each other, to name for each sequence three individuals who may have generated the sequence. Whereas the recognition rate was rather poor for other participants' clapping, it was much higher for the participants' own clapping. Repp attributed this finding to individual differences in hand configuration when clapping, which affected the acoustic properties of individual claps. However, because the participants had clapped at a self-selected rate, individual differences in tempo and timing may also have contributed to self-recognition.

In order to address this issue, we used a design in which each participant listened to their own and just one other participant's clapping sequences. These paired individuals were not known to each other. They did not have the opportunity to listen to all the recordings before making the judgments. Furthermore, in one experimental group ("full information") the recordings were reproduced in their original format, whereas in a second experimental group ("temporal information"), the claps were replaced by pure tones having a constant duration and amplitude, so that only the temporal structure of the clapping sequences was retained. To the extent that self-recognition relies on acoustic properties, the replacement of the claps by pure tones should impair self-recognition. If self-recognition relies only on temporal properties, however, this manipulation should not matter.

Like Repp (1987), we allowed the participants to generate the clapping sequences at a self-selected rate so that inter-individual differences in general tempo were preserved. This could have been crucial because individuals differ in the tempo they spontaneously choose. Thus, Rimoldi (1951) asked participants to carry out a large sample of movements in a "natural, congenial way." Using factor analysis, he then distinguished between the tempo of large movements (e.g., moving arms in parallel), the tempo of small movements (e.g., tapping with the right index finger), and the tempo of other movements (e.g., drawing circles). Repeating the tests 2 to 4 weeks later, he found retest reliability coefficients of

.71 to .90 for large movements and of .87 to .92 for small movements. The assumption that the preferred tempo is important is also supported by Fraisse (1982). He reported that participants, instructed to generate an irregular temporal sequence, usually only generate two different intervals. Either the durations of two adjacent intervals tend to be the same, or they tend to have a ratio of roughly 1:2 (cf. also Povel, 1981). This distinction between long and short intervals seems to be sharpest at the tempo participants spontaneously choose.

In contrast to Repp (1987), we also varied the complexity of the sequences. In particular, the ratios of temporal intervals could be 1:1, 1:2, 1:3, or 2:3 (Fig. 1). Thus, some sequences were easier to generate than other sequences, which could provide additional opportunities for inter-individual differences in the accuracy of their reproduction to occur. In addition, more complex sequences provide more possibilities to group individual claps in different ways (cf. also Fraisse, 1982).

Moreover, we asked our participants to reproduce the temporal patterns from the beginning of the test session once more, this time using a different movement. By doing this we tried to assess the temporal consistency of different movement parameters that might contribute to an authorship effect—temporal consistency is a prerequisite for an off-line authorship effect, given that the effect is not mediated by episodic memory processes.

Finally, we instructed the participants in the test session to press a key in synchrony with the auditory sequences. Because participants may be better at synchronizing with self-generated than with other auditory sequences, their synchronization performance may provide additional cues for self-recognition. In summary then, we increased the temporal information that could be used for self-recognition in comparison to Repp's study.

Method

Participants

Thirty-two participants took part in this experiment, 8 of whom were men. They ranged in age from 16 to 31 years. Three participants were left-handed. All

a. 1:1:1:1:1:1:1:1   b. 2:2:2:2   c. 2:1:2:1:2   d. 3:3:2

e. 3:1:2:2   f. 2:1:3:2   g. 2:1:1:1:2   h. 2:4:2
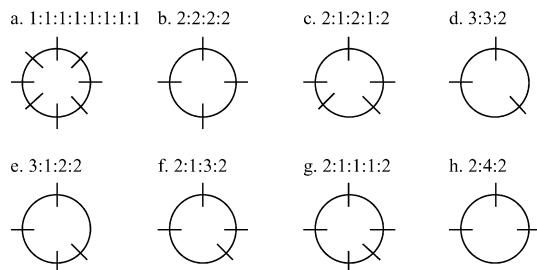
Fig. 1 Patterns used in Experiments 1–3. Alphabetical order corresponds to the order of presentation. Numbers indicate the ratios of the individual durations

participants received payment for their participation. Half of them were randomly assigned to the full information group and half to the temporal information group.

Apparatus and stimuli

In the recording session, the participants sat in front of a desk, with their elbows resting on the desk top. The clapping sequences were recorded with a microphone located approximately 10 cm above the participants' hands. For eight participants, a SONY ECM-260F microphone and a RIM amplifier was used. For the remaining participants, an Apple PlainTalk microphone was used. The clapping patterns to be produced were presented on an Apple 17" monitor, which was located to the participants' right. Fig. 1 illustrates the format in which different clapping patterns were presented. In addition, numbers were added inside and outside each circle. Numbers outside the circle indicated beat positions within a period. Numbers inside the circle indicated the number of beats to the next clap. The start of the recording was indicated by the emergence of an exclamation mark, displayed in the middle of the circle. At the end of the recording, the screen turned white. During recording, the amplitude of the acoustic signal was measured every 10 ms. This measurement was then used to determine the times of the maximum amplitudes of the claps, at which the onsets of the pure tones were presented in the temporal information group.

All in all, we used eight different clapping sequences (cf. Fig. 1). In order to generate the first sequence, participants clapped regularly at their most comfortable rate. For the second sequence, they omitted every other clap. Whereas the durations of the individual intervals needed to be changed, the underlying beat and, therefore, the general tempo could be maintained. The same was true for the remaining sequences. They differed from the preceding sequences in that they consisted of intervals of different durations, in ratios of 1:2, 1:3, or 2:3.

The stimuli in the test session were presented via audio-techniqua ATH-PRO6 headphones. The amplitude of the acoustic recordings was adjusted to a comfortable level. The recordings were adjusted so that the first clap occurred after 500 ms of silence. The same was true for the first tone in a tone sequence. The tones were generated by a square-wave synthesizer. They lasted for 10 ms, and their frequency was 261.625 Hz—corresponding to middle C. All participants for whom the SONY microphone had been used were in the full information group.

The synchronized responses were recorded with a PsyScope button box. The recognition judgments were recorded as follows: At the end of a trial, a horizontal line with a length of 15.39 cm was displayed on the screen. At the left end of the line, which was marked by a small vertical line, the German word *selbst* (self) was displayed. At the right end of the line, also marked by a

small vertical line, the word *fremd* (other) was displayed. Using the computer mouse, the participants placed a vertical line between these two end points. By pressing the mouse button, the position of the vertical line was recorded.

Procedure

The experiment consisted of two sessions separated by at least 1 week. In the recording session, the eight clapping sequences were recorded. At the beginning of a trial, the sequence to be produced was displayed on the screen (Fig. 1). The participants were instructed to generate the sequence repeatedly without interruption. Accordingly, the line at the 12 o'clock position indicated both the last clap of the preceding period and the first clap of the next period. Participants could use any aid that helped them to generate the instructed temporal pattern. For instance, they were allowed to count silently. The only constraint was that speech, foot tapping and the like should not be audible. Importantly, participants were encouraged to choose a tempo that felt most natural to them and to maintain this tempo during the whole recording session. However, they were not informed that they would be asked to identify their own clapping in a later session. The different sequences were recorded in the order indicated in Fig. 1. Participants practiced each sequence until they were able to generate it correctly. Only then did the actual recording take place. It started with a countdown of the numbers 3, 2, and 1 centrally displayed on the screen. Each number was displayed for 500 ms. Afterwards the visual display of the sequence was shown and participants started to reproduce the intervals starting from the 12 o'clock position. The recording lasted 30 s. The session ended with a brief interview about the strategies the participants had used to generate the sequences. The session lasted 20 to 40 min.

In the test session, two conditions were realized. Half of the participants heard original recordings of clapping (full information group). The other half heard tone sequences, which reproduced the temporal succession of the maximum amplitudes of the claps (temporal information group). Two paired participants heard the same auditory events in the same order. Half of the sequences reproduced self-generated clapping and the other half reproduced other-generated clapping. Each trial started with the presentation of the corresponding static visual display of a sequence, which remained visible during the whole trial. Participants were asked to practice clapping the sequence again. Then they pressed a key, and after the countdown described above the respective clapping or tone sequence was played.

The participants were instructed to press a key in synchrony with each clap. The participants in the full information group were instructed to synchronize each key press with the maximum loudness of each clap. The participants in the temporal information group were told to synchronize each key press with the occurrence of the tone. After each presentation, the participants used the continuous scale described above to judge whether or not they had generated the auditory sequence. They were instructed that setting the vertical line to the left of the midpoint would indicate a "self" judgment and setting the vertical line to the right of the midpoint would indicate an "other" judgment. They were also instructed that the distance of the vertical line from the midpoint would indicate their confidence in their judgment.

Each participant received three blocks, each containing the same 16 sequences. The sequence order within a block was pseudo-random, to make sure that self-generated and other-generated clapping sequences appeared roughly equally often at the beginning and at the end of the block. At the beginning of the test session, the participants repeated the task of the recording session. With the exception that they now used key presses instead of claps, all other details remained constant. The test session ended with a brief interview about the strategies used for self-recognition. The session lasted about 60 min.

Data analysis

The recognition judgments were measured on a scale with a range between 0 and 1, using a resolution of .01. From this scale, we derived a dichotomous measure of self-recognition in order to obtain independent measures of discrimination accuracy and response bias. The response bias measure can be used to identify those stimulus parameters that prompt participants to believe that they had generated a sequence. The measure of discrimination accuracy can be used to identify those stimulus parameters that enabled participants to veridically identify their own clapping sequences. Additionally, we obtained a measure of judgmental confidence by determining the absolute difference between the actual judgment value and the criterion value—a value higher than .50 was coded as indicating self-attribution. Because the source (self or other) of the clapping sequence was also a dichotomous variable, the number of hits (H), false alarms (FA), misses (M), and correct rejections (CR) could be determined. In analogy with signal detection methodology, a non-parametric discriminability index $d*$ and a non-parametric response bias index $\beta*$ was calculated (before, we collapsed the data across those variables that were not used in the subsequent analysis of the two measures):

$$d* = \frac{H - FA}{H + M} \text{ and } \beta* = \frac{H + FA}{H + FA + M + CR}$$

The discriminability index $d*$ can vary between -1 and 1. Positive values indicate correct self-identification. The response bias index $\beta*$ can vary between 0 and 1. Values higher than .50 indicate that sequences tended to be judged as self-generated and values lower than .50 indicate that sequences tended to be judged as

other-generated. In analogy with correlations, a Fisher-z-transformation was applied to these measures before calculation.

Results

Preliminary analyses of the sequence variable (a, b, c, d, e, f, g, h) did not show significant effects on $d*$, $F(7, 217) = .37$ ($p = 0.92$), so we collapsed the data across this variable. We also collapsed the data across the block variable (1, 2, 3) because it did not show a significant effect on $d*$, $F(2, 60) = 1.92$ ($p = .16$) either.

The discrimination measure $d*$ was then tested against the chance level of .00 for the full information group ($M = .55$; $s = .47$) and for the temporal information group ($M = .50$; $s = .50$) separately. Both t-tests were significant [full information group: $t(15) = 4.83$ ($p < .01$); temporal information group: $t(15) = 4.00$ ($p < .01$)]. The groups did not differ significantly: $t(30) = .34$ ($p = .74$). Individual $d*$-values ranged from $-.29$ to .92. Given the high inter-individual variability, a 95%-confidence interval was calculated for the chance value of .00; 24 out of 32 participants exceeded its upper limit of .19.

The response bias measure $\beta*$ was tested against the chance level of .50 for the full information group ($M = .52$; $s = .18$) and for the temporal information group ($M = .48$; $s = .12$) separately. None of the t-tests showed a significant effect [full information group: $t(15) = 1.60$ ($p = .30$); temporal information group: $t(15) = .82$ ($p = .43$)]. Moreover, the groups did not differ significantly: $t(30) = .90$ ($p = .38$).

Discussion

First of all, Repp's (1987) observation that the participants can recognize their own clapping was confirmed. Moreover, this finding was not restricted to temporally regular sequences. Importantly, however, the temporal information contained in the clapping sequence turned out to be sufficient for the recognition of one's own clapping.

The fact that the acoustic properties of the claps did not provide additional cues for self-recognition does not necessarily mean that they played no role. Not only were the sequences used in the present experiment more complex than Repp's (1987) sequences, but the instructions also directed the participants' attention to temporal properties.[1] For this reason, the temporal information may have been more salient than the acoustic information. Also, it is conceivable that paired participants who differed in the temporal properties of

their claps also differed in the acoustic properties of their claps. Further research is needed to settle this issue.

As mentioned above, self-recognition performance did not differ significantly between different sequences. This was surprising because the sequences differed significantly with regard to the accuracy of their reproduction, $F(7, 217) = 7.53$ ($p < .01$). However, the accuracy of sequence reproduction did not significantly predict inter-individual differences in self-recognition performance. In contrast, the general tempo, determined over all sequences, significantly predicted inter-individual differences in self-recognition performance.[2]

Experiment 2

The second experiment aimed at assessing the relative importance of the inter-individual differences in general tempo for the recognition of self-generated clapping sequences. For this purpose, all participants listened to pure tone sequences reproducing the temporal pattern of clapping. This allowed us to vary the general tempo orthogonally to the other temporal information. In particular, each participant's own sequences and the other participant's sequences in each pair were presented both in the participant's own tempo and in the other participant's tempo. If the participants exclusively base their recognition judgments on inter-individual differences in general tempo, they should attribute their own sequences to themselves when they are presented in their own tempo, but to the other participant when they are presented in the other tempo. To the degree that they base their judgments on other temporal information, however, the tempo manipulation should not matter.

The original reproductions of the claps in Experiment 1 provided information about pitch, duration, and amplitude in addition to temporal information. Variations in these parameters may indicate an accentuation of individual claps, based on their position in the sequence. In other words, these parameters may indicate a grouping of sequence events (cf. Fraisse, 1982). In the present experiment, half of the participants listened to

---

[1] We are grateful to Bruno Repp for suggesting this interpretation. In Repp's (1987) study, the participants were not asked to maintain a constant tempo in the recording session. In addition, they were not asked to accompany the recorded claps with synchronization responses in the test session.

[2] We determined the general tempo of a sequence by the average of all complete sequence periods. The accuracy of sequence reproduction was determined by the ratios of adjacent intervals. These observed ratios were then compared with the expected ratios, as indicated in Fig. 1, by calculating the differences, squaring them, adding up the squared differences, and deriving the square root of the sum. This variable error has the advantage that it is logically independent of the differences in general tempo. Next, we assessed the reliability of these inter-individual difference measures. Overall correlations of these measures between the two sessions as well as individual comparisons for each participant pair showed that the general tempo, but not the accuracy, of sequence reproduction was a reliable inter-individual difference measure. A standard regression analysis of the self-recognition measure revealed only general tempo to be a significant predictor. Pairs of participants who showed large differences in general tempo were more likely to recognize their own clapping sequences than pairs of participants who did not show large differences in general tempo.

the same tone sequences as the temporal information group in Experiment 1. The other half (the "accent information" group) listened to a sequence of tones that retained the information about the durations and maximum amplitudes of the claps. An example of the time-amplitude functions of a clap and its reproduction is shown in Fig. 2.

Method

Participants

Thirty-two new participants took part in this experiment, 8 of whom were men. They ranged in age from 18 to 27 years. One participant was left-handed. All participants received payment for their participation. The participants were randomly assigned to the temporal information group and the accent information group.

Apparatus and stimuli

The same apparatus was used as in the previous experiment. This time, however, all recordings were done with an Apple PlainTalk microphone.

The stimuli in the recording session were identical to the ones in the previous experiment. In the test session, tones were either presented at the same tempo with which the claps had been generated, or they were presented at the general tempo of the paired participant. As a consequence, the durations of the transformed reproductions could vary. If the generated sequences were reproduced at a faster tempo, they lasted less than 30 s. If they were reproduced at a slower tempo, they were truncated at 30 s. In the temporal information group, the tones lasted for 10 ms, and the amplitude was about 36% of the maximum amplitude (Fig. 2). In the accent information group, amplitudes of the tones reproduced the maximum amplitude of the claps, and the durations

of the tones reproduced the durations between the occurrence of the maximum amplitude and the subsequent drop of the amplitude below 10% of the maximum amplitude (Fig. 2).

Procedure

There were two sessions. The recording session was identical to that in the previous experiment. The test session differed from the test session of the previous experiment in the following way. Most importantly, the eight self-generated and the eight other sequences were presented once in the original tempo and once in the tempo of the paired participant. This resulted in two blocks of 16 trials each, which were presented in pseudo-random order. In addition, there were the two presentation conditions described above. Once again, the participants started the test session with a repetition of the recording session, with the difference that they used key presses instead of claps. At the end of the test session, the participants were asked whether they had noticed that some of the sequences were not reproduced in the original tempo.

Data analysis

The same analyses were carried out as in the previous experiment.

Results

We once more collapsed data across the sequence variable, because it did not significantly affect d^*, $F(7, 168) = .69$ ($p = .69$). Separate 2×2×2-ANOVAs with the between-subject variable group and the within-subject variables tempo condition (untransformed, transformed) and block (1, 2) were then calculated for d^* and β^*.

The ANOVA showed a significant main effect of tempo condition, $F(1, 30) = 14.56$ ($p < .01$). In addition, the block effect approached significance, $F(1, 30) =$

Fig. 2 Time-amplitude functions of the auditory events reproduced in Experiments 1–3. The *full line* indicates the original recording, reproduced in the full information group in Experiment 1. The *dashed line* indicates the reproduction in the accent information group of Experiment 2. The *dotted line* indicates the reproductions in the remaining conditions. On the abscissa, time is denoted. On the ordinate, the amplitude is denoted relative to its maximum
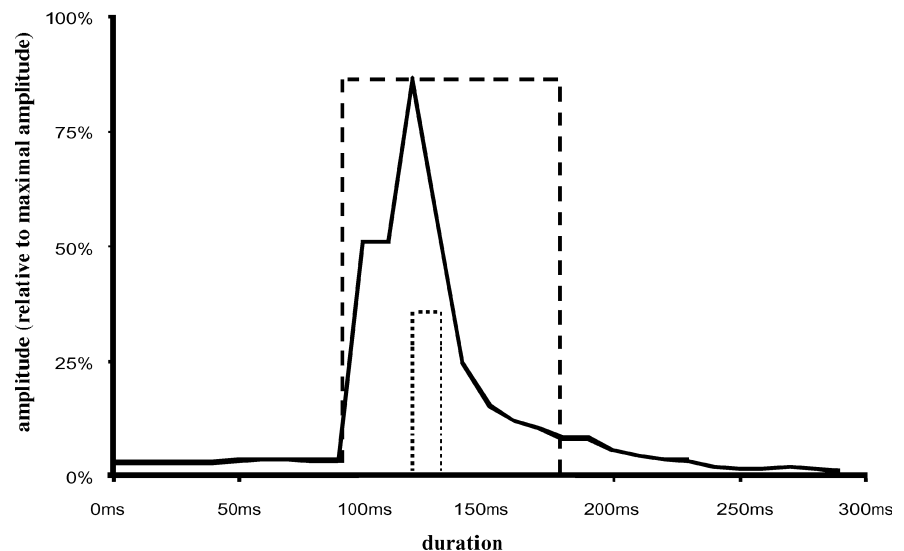
3.84 ($p = .06$). Whereas $d*$ was significantly above chance for the untransformed sequences, first block: $M = .42$; $s = .50$; $t(31) = 4.75$ ($p < .01$); second block: $M = .36$; $s = .47$; $t(31) = 4.28$ ($p < .01$), it did not significantly differ from zero for the transformed sequences, first block: $M = .02$; $s = .43$; $t(31) = .30$ ($p = .77$); second block: $M = -.09$; $s = .44$; $t(31) = -1.19$; ($p < .24$).

The analysis of the response bias measure only showed a non-significant interaction between block and tempo condition, $F(1, 30) = 3.42$ ($p = .07$). Whereas $ß*$ deviated from chance for the transformed sequences in the second block, $M = .42$; $s = .20$; $t(31) = -2.09$ ($p = 0.04$), it did not differ significantly from .50 in the other conditions ($M = .49$ in all other conditions).

Discussion

The main result of the second experiment is that the tempo manipulation disturbed the self-recognition performance, which indicates that the participants used the general tempo as a cue for self-recognition. Whereas the self-recognition performance for the untransformed sequences was similar to that in Experiment 1, it dropped to chance for the transformed sequences. If the general tempo had been the only cue for self-recognition, the discrimination measure should just have reversed its sign for the transformed sequences. This was not the case, however.

The additional accent information provided to half of the participants did not significantly enhance self-recognition. This may be attributed to the fact that the participants in this experiment did not differ consistently in their use of accents, as given by the tone durations and maximum amplitudes.

It should be noted that only a quarter of all presented sequences reproduced exactly a temporal pattern that had actually been generated by a given participant during the recording session. Nevertheless, participants judged about half of the presented sequences as self-generated. Only in the second block, this tendency declined as indicated by the reduction in the response bias measure (.42). This indicates that participants based their judgments on more than one cue. If they had only used the relative temporal information, which was preserved in untransformed and transformed sequences, or if they had only used the general tempo, the response bias should have amounted to .50. Only by the use of multiple cues could they become aware of the tempo manipulation and develop a bias towards judging sequences as generated by someone else.

Experiment 3

The previous experiment confirmed that inter-individual differences in general tempo provide one cue for self-recognition. Moreover, it showed that additional cues

may be available. In particular, it is conceivable that synchronization performance contributed to self-recognition as well.[3]

In order to test this hypothesis, we replicated the temporal information condition of Experiment 1 with the difference that the participants were not asked to synchronize with the auditory sequences. To the extent that the self-recognition performance relies on temporal cues that are not bound to concurrent synchronization, it may even be predicted that self-recognition would improve in this condition, because the synchronization task may have diverted attentional resources away from the self-recognition task.

To the extent that self-recognition performance relies on cues that are bound to concurrent synchronization, self-recognition should deteriorate in this experiment. However, it is conceivable that, contrary to instructions, participants spontaneously engage in some kind of covert synchronization in order to discriminate self-generated sequences from other-generated sequences. To prevent this from occurring, we introduced a different secondary task in a second ("dual-task") condition. This task consisted of continuously drawing circles, at either a fast or a slow tempo. In the first condition ("single-task"), participants just listened to the temporal sequences.

If the self-recognition judgments are less accurate in the dual-task condition than in the single-task condition, attentional resources seem to be required for self-recognition. If the instruction to draw fast or slow interacts with self-recognition performance, a representation of general tempo seems to be needed for action generation.

The present experiment differed from the previous ones in one more detail. The participants no longer passed a practice phase at the beginning of the test session. Arguably, this reproduction of the temporal sequences could have reminded them of their temporal performance. To the extent that this tapping performance reliably reproduces the temporal characteristics of their previous clapping performance, participants could have compared the perceived clapping sequences with the tapping sequences they had just generated. Omitting this tapping task at the beginning of the test session this time, participants could no longer use this

---

[3]In contrast to Experiment 1, Experiment 2 revealed a significant effect of authorship on synchronization performance. Participants were less variable when synchronizing with self-generated sequences than when synchronizing with other sequences. Moreover, the more variable the participant's synchronization performance was with regard to the other participant's sequences than with regard to his or her own sequences, the more often the self-recognition judgments were correct. The fact that the analysis of the synchronization performance did not reveal more significant effects of authorship may be rooted in the fact that the overall synchronization performance was rather variable and error-prone. This fact was especially challenging for the pre-analysis of the synchronization data. Nonetheless, the analysis of the synchronization data reliably showed significant block effects for the constant and the variable error of synchronization, decreasing with increasing practice.

proxy of their original performance. As a consequence, self-recognition may deteriorate.

Method

Participants

Sixteen new participants took part in this experiment, 6 of whom were men. They ranged in age from 21 to 36 years. Three participants were left-handed. All participants received payment for their participation. Due to a coding error, one pair of participants had to be omitted from analysis.

Apparatus and stimuli

The same apparatus was used as in the previous experiment. In the recording session, the drawing movements were recorded with a WACOM 2.5.5-D digitizer tablet with a size of 45.72×30.48 cm. Its spatial resolution was $22,860 \times 15,240$ pixels; the sampling rate was 100 Hz. The mapping between the graphic tablet and the screen was 1:1. The room light was dimmed. The background of the computer monitor was set to black.

The drawing movements were displayed as a black dot with a diameter of about .38 cm on the screen. The hand movements were hidden from view by a cover mounted on the writing pad. The black dot moved within a white ring whose inner border had a diameter of 7.69 cm and whose outer border had a diameter of 15.38 cm. Within the inner border, another circle with a diameter of 5 cm indicated the appropriate sequence. In the middle of the circle, the movement instructions were displayed, namely the German words *schnell* (fast), *langsam* (slow), or *nicht bewegen* (do not move). The recognition judgments were recorded in the same way as in the previous experiments, with the exception that this time the participants used the pen of the graphic tablet to move a vertical line on the screen.

Procedure

There were two sessions, the first of which was identical to the first session of the previous experiments. The second session started with a practice block that familiarized participants with the drawing utensils as well as with the course of each trial. Participants started the trial by moving the dot to a fixed starting position, indicated by a second black dot at the ring's twelve o'clock position. When the dot had arrived there, the recording was started. It always lasted for 30 s. Participants started to draw circles at their most comfortable tempo. In the following trials, they drew at a markedly faster and, eventually, at a markedly slower tempo. After the practice trials, participants completed two dual-task blocks (D) and two single-task blocks (S), presented in alternating order (D-S-D-S for half of the participants,

S-D-S-D for the other half). Each block consisted of 16 trials in which the participants listened to a tone sequence reproducing either their own or another participant's clapping. At the end of each trial, they indicated whether or not the sequence was self-generated. In the dual-task blocks, they drew circles while listening to the tone sequences. The movement instruction for each trial followed the same pseudo-random order for all participants.

Data analysis

The same dependent measures of self-recognition were used as in the previous experiments. In addition, we analyzed the average drawing velocity in order to verify that all participants followed the movement instructions.

Results

Preliminary analyses of the self-recognition judgments showed that the sequence variable did not significantly affect d^*, $F(7, 91) = .91$ ($p = .50$), so that the data were once again collapsed across this variable.

Separate 2×2-ANOVAs of the judgment measures with the within-subject variables of condition (dual-task, single-task) and block (1, 2) revealed no significant effects. Separate t-tests showed that d^* was significantly above chance for the dual-task condition, $M = .35$; $s = .50$; $t(13) = 2.42$ ($p = .03$), as well as for the single-task condition, $M = .28$; $s = .40$; $t(13) = 2.43$ ($p = .03$). None of the $ß^*$-values differed significantly from the chance level of .50 (dual-task condition: $M = .49$; $s = .13$; single-task condition: $M = .50$; $s = .12$).

Focusing on the judgments in the dual-task trials, we analyzed whether the velocity instruction (fast, slow) significantly affected d^* (fast: $M = .43$; $s = .49$; slow: $M = .36$; $s = .71$) or $ß^*$ (fast: $M = .47$; $s = .14$; slow: $M = .53$; $s = .16$). Neither difference was significant [d^*: $t(13) = .45$ ($p = .66$); $ß^*$: $t(13) = 1.66$ ($p = .12$)].

Subsequently, we classified the trials with regard to the congruency between the inter-individual differences in general tempo and the velocity instruction—e.g., a trial was classified as congruent if the faster participant in a pair of participants followed the fast instructions and if the slower participant followed the slow instructions. No significant differences between congruent trials and incongruent trials could be observed with regard to d^*, congruent: $M = .32$; $s = .52$; incongruent: $M = .48$; $s = .69$; $t(13) = 1.11$ ($p = .29$) or $ß^*$, congruent: $M = .49$; $s = .15$; incongruent: $M = .63$; $s = .60$; $t(13) = 1.12$ ($p = .08$).

In addition, we calculated the product-moment correlations between the general tempo of clapping and the average velocity of drawing (fast: $r = .32$; slow: $r = -.13$), neither of which was significant [fast: $t(10) = 1.08$ ($p = .31$); slow: $t(10) = -.42$ ($p = .68$)].

Discussion

The overall self-recognition performance was above chance, indicating that self-recognition does not hinge on the concurrent synchronization with the auditory sequences. Self-recognition performance was somewhat less accurate than in Experiment 1, but the difference did not reach significance: $t(28) = 1.67$ ($p = .11$).

The fact that the involvement of the participants in a different secondary task did not significantly affect the results shows that concurrent synchronization is not necessary for self-recognition to occur. Thus, self-recognition performance cannot be attributed exclusively to concurrent motor activation. It seems that participants can base their self-recognition judgments on more abstract features. However, despite our procedures, participants could have tapped with their feet in synchrony or counted silently. However, because the velocity of the drawing movement was not influenced by the tempo of the auditory sequences, it has to be assumed that the temporal control of drawing movements partly differs from the temporal control of the tapping movements (cf. also Zelaznik, Spencer, & Doffin, 2000). This may explain why the drawing task did not interfere with the self-recognition task.

General discussion

The present study extends previous evidence of off-line authorship effects in action perception. The previous evidence stems from tasks in which spatially defined stimuli evoked visual action effects. In the present study, temporally defined stimuli evoked auditory action effects. Thus, we conclude that off-line authorship effects are not restricted to the visual modality. Moreover, they need not be mediated by spatial representations.

According to the common coding approach to perception and action planning, action representations consist of individual features, such as left or right or high pitch or low pitch. Moreover, it has been assumed that the activation of a single feature—e.g., by the presentation of a tone with high or low pitch—automatically activates the whole action representation. As a consequence, a left or right response is more likely to occur (cf. Hommel et al., 2001). Following this line of thought, it is conceivable that a slow tempo is associated with rhythmical grouping A (e.g., two long intervals followed by one short interval) and a fast tempo is associated with rhythmical grouping B (e.g., four long intervals). If the rhythmical grouping A occurred at a faster tempo than rhythmical grouping B, the sequence would not be recognized as self-generated. According to the theory, these feature bundles are assumed to be instantiated by motor processes that are not part of the action representation itself. This is an important difference from other conceptualizations of perception-action interactions.

Off-line authorship effects in the temporal domain could also result from the operation of a modular timing device (cf. Keele et al., 1985). This timing device presumably works best for a given range of durations (cf. Fraisse, 1982; Parncutt, 1994). It is also conceivable that individuals differ with regard to their optimal tempo (cf. Rimoldi, 1951) or with regard to their variability being associated with a given tempo (cf. Smoll, 1975; Smoll & Schutz, 1978). Moreover, the same assumptions may hold true for an internal simulation of the respective action (cf. Decety & Jeannerod, 1996; Sirigu et al., 1996).

In order to judge whether or not a perceived clapping sequence has been self-generated, participants may internally simulate the action and compare the result of the simulation with the perceived action effect. Thus, they may notice that the perceived clapping sequence has been generated in a faster tempo than they would choose to generate it. Or, they may try to simulate the action in the same tempo as the perceived action and notice that their simulation is more variable or error-prone than it used to be.

In a synchronization task, participants have to adjust their own tempo to the tempo of the perceived clapping sequence. If they are just a bit too fast or too slow, errors will inevitably accumulate. The resulting error corrections may manifest themselves in an increased variability in the synchronization performance (cf. Vorberg & Wing, 1996). In Experiment 2, we found an increased variability when the participants synchronized with a clapping sequence that was not generated by themselves in comparison to a clapping sequence that was generated by themselves. Altogether, however, the synchronization data did not provide compelling evidence for the assumption that we synchronize with a self-generated clapping sequence in a different way than we synchronize with another clapping sequence.

This null effect may be partly due to difficulties in the analysis of the synchronization data—i.e., the overall synchronization performance was rather error-prone. Because the omission of the synchronization task as well as its replacement by a different motor task in Experiment 3 did not prevent self-recognition, we are inclined to conclude, however, that the processes involved in self-recognition are partly different from the motor processes that are involved in the generation of the action effect.

Certainly, more research has to be carried out in order to decide this issue. One possibility may be to look at expert performance. It is known that rhythmical expertise, at least for longer intervals, is accompanied by lower intra-individual variability in synchronization as well as in continuation tapping tasks (cf. Gérard & Rosenfeld, 1995). Thus, potential inter-individual differences in synchronization with self-generated and other temporal sequences are more likely to turn out to be significant. Moreover, rhythmical expertise seems to allow individuals to integrate musical events over longer time spans (cf. Drake, Penel, & Bigand, 2000). In other words, it becomes more likely that a given rhythmical

grouping as such is represented at all (cf. also Smith, 1982). Then, it may be possible to directly compare the influence of different stimulus factors on self-recognition as well as on motor performance being measured at the same time.

Finally, we would like to address a point that most naturally comes to mind when conceding that our off-line authorship effects only turned up in a self-recognition task. Thus, one may wonder in which way the measured effects are different from memory effects. To be clear, action representations, as conceived in the common coding approach to perception and action planning, are acquired by experience, so that they can be conceived of as memory representations. These action representations differ from episodic memory representations, however, by the fact that they are used each time the respective action is generated. They do not presuppose a conscious recollection of the episode of the action generation. However, they also do not exclude this possibility.

Our data do not allow us to unambiguously distinguish between these possibilities. We do not know to what extent the correct recognition of a clapping sequence as self-generated was associated with the remembrance of a particular performance during the recording session. However, we do know that the stimulus parameter that proved to be reliable over two test sessions—i.e., the general tempo—was also the stimulus parameter that primarily influenced self-recognition performance. In the present experiments this temporal cue was more important to action generation than non-temporal cues, such as the sounds of clapping. To our knowledge, an episodic memory explanation gives no indication why some cues turned out to be more important for self-recognition than others. Therefore, the present data seem to be consistent with the common coding approach (cf. Hommel et al., 2001). How this framework can be reconciled with the assumption of a conscious mediation of recognition effects or with an episodic memory explanation remains a question that may stimulate future research.

References

Decety, J., & Jeannerod, M. (1996). Mentally simulated movements in virtual reality: Does Fitt's law hold in motor imagery. *Behavioural Brain Research, 72*, 127–134.
Drake, C., Penel, A., & Bigand, E. (2000). Tapping in time with mechanically and expressively performed music. *Music Perception, 18*, 1–23.
Flach, R., Knoblich, G., & Prinz, W. (2003). Off-line authorship effects in action perception. *Brain & Cognition, 53*, 503–513.
Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 149–180). New York, Academic Press.

Gérard, C., & Rosenfeld, M. (1995). Pratique musicale et régulations temporelles [Musical practice and temporal control]. *L'Année psychologique, 95*, 571–591.
Greenwald, A. G. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideo-motor mechanism. *Psychological Review, 77*, 73–99.
Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC). A framework for perception and action. *Behavioral & Brain Sciences, 24*, 849–937.
Ivry, R. B., & Hazeltine, E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 3–18.
Ivry, R. B., & Keele, S. W. (1989). Timing functions of the cerebellum. *Journal of Cognitive Neuroscience, 1*, 136–152.
Keele, S. W., Pokorny, R. A., Corcos, D. M., & Ivry, R. (1985). Do perception and motor production share common timing mechanisms: A correlational analysis. *Acta Psychologica, 60*, 173–191.
Knoblich, G., & Prinz, W. (2001). Recognition of self-generated actions from kinematic displays of drawing. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 456–465.
Knoblich, G., Seigerschmidt, E., Flach, R., & Prinz, W. (2002). Authorship effects in the prediction of handwriting strokes: Evidence for action simulation during action perception. *Quarterly Journal of Experimental Psychology, 55*, 1027–1046.
Parncutt, R. (1994). A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception, 11*, 409–464.
Povel, D.-J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 3–18.
Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*, 129–154.
Repp, B. (1987). The sound of two hands clapping: An exploratory study. *Journal of the Acoustical Society of America, 81*, 1100–1109.
Rimoldi, H. J. A. (1951). Personal tempo. *Journal of Abnormal and Social Psychology, 46*, 283–303.
Sirigu, A., Duhamel, J.-R., Cohen, L., Pillon, B., Dubois, B., & Agid, Y. (1996). The mental representation of hand movements after parietal damage. *Science, 273*, 1564–1568.
Smith, J. (1982). Reproduction and representation of musical rhythms: The effects of musical skill. In D. Rogers & J. A. Sloboda (Eds.), *The acquisition of symbolic skills* (pp. 273–282). New York, Plenum.
Smoll, F. L. (1975). Preferred tempo of motor performance: Individual differences in within-individual variability. *Journal of Motor Behavior, 7*, 259–263.
Smoll, F. L., & Schutz, R. W. (1978). Relationships among measures of preferred tempo and motor rhythm. *Perceptual and Motor Skills, 46*, 883–894.
Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception, 19*, 705–743.
Treisman, M., Faulkner, A., & Naish, P. L. N. (1992). On the relation between time perception and the timing of motor action: Evidence for a temporal oscillator controlling the timing of movement. *Quarterly Journal of Experimental Psychology, 45A*, 255–263.
Vorberg, D., & Wing, A. (1996). Modeling variability and dependence in timing. In H. Heuer & S. W. Keele (Eds.), *Handbook of Perception and Action* (vol. 2, pp. 181–262). San Diego, CA: Academic Press.
Zelaznik, H. N., Spencer, R. M., & Doffin, J. G. (2000). Temporal precision in tapping and circle drawing movements at preferred rates is not correlated: Further evidence against timing as a general-purpose ability. *Journal of Motor Behavior, 32*, 193–199.