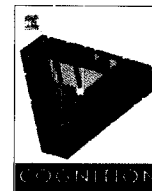




ELSEVIER

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Inferring agency from sound

Günther Knoblich^a, Bruno H. Repp^{a,b,*}

^aDonders Institute for Brain, Cognition, and Behavior, Centre for Cognition (DCC), Radboud University Nijmegen, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

^bHaskins Laboratories, 300 George Street, New Haven, CT 06511-6624, USA

ARTICLE INFO

Article history:

Received 10 March 2008

Revised 19 February 2009

Accepted 20 February 2009

Keywords:

Agency

Agency cues

Self

Synchronization

Rhythm

Tapping

ABSTRACT

In three experiments we investigated how people determine whether or not they are in control of sounds they hear. The sounds were either triggered by participants' taps or controlled by a computer. The task was to distinguish between self-control and external control during active tapping, and during passive listening to a playback of the sounds recorded during the active condition. Experiment 1 required detection of a change in control mode within trials. Experiments 2 and 3 introduced a simple rhythm reproduction task that requires discrimination of control modes between trials. The results demonstrate that both sensorimotor cues and perceptual cues are used to infer agency. In addition, there may be further influences of cognitive expectation and/or multimodal integration. In accordance with hierarchical models of intention [e.g., Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107, 179–217] this suggests that the sense of agency is not situated on one specific level of action control but subject to multiple influences.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Consider the following prototypical scene that may occur in any number of crime movies: An unsuspecting man walks along a dark, lonely street. The shabby houses along the street reflect a loud echo of what appear to be his footsteps. Behind him is a dark figure walking at the same pace, closing up on him. The viewer hopes that the poor victim will notice that there is another, malicious, person in the scene. And indeed, at some point the victim turns around, suspecting he is not alone. But it's too late. The screen turns dark.

The example illustrates that, although it is crucial to determine whether one is causing the perceptual events one perceives, it is not always easy. This is the problem of agency. How do I know whether I am causing the rhythmic sounds of footsteps? When do I begin to suspect that another actor is causing some of these sounds? Do cognitive expectations, such as my belief that I am walking alone, affect agency? In this article we report results from three experiments obtained with a new experimental paradigm that allowed us to determine how people derive agency from sound.

1.1. Previous research

Previous research on agency has explored a number of potential cues to agency, and several different cognitive mechanisms for agency have been proposed. First, there is Wegner's (2002) provocative claim that our experience of having a conscious will is an illusion. The implication of his framework is that one will experience agency for an event when (1) one's action preceded the event (priority), (2) one's action is consistent with the event (consistency), and (3) when there are no obvious alternative causes for the event. Wegner and his colleagues demonstrated in a variety of studies that people develop feelings of illusory control over events when these three

* Corresponding author. Address: Haskins Laboratories, 300 George Street, New Haven, CT 06511-6624, USA. Tel.: +1 203 865 6163x236; fax: +1 203 865 8963.

E-mail addresses: Knoblich@donders.ru.nl (G. Knoblich), repp@haskins.yale.edu (B.H. Repp).

conditions are given (Aarts, Custers, & Wegner, 2005; Wegner, Fuller, & Sparrow, 2003; Wegner & Wheatley, 1999).

Although Wegner's claims have led to a renewed interest in agency research, his account seems a bit pessimistic with regard to the existence of cognitive and brain mechanisms that actually provide accurate information about whether one has caused perceived events or not. Pacherie's (2008) conceptual framework for the phenomenology of action seems to provide a more balanced account (cf. also Proust, 2000). She distinguishes between three hierarchically ordered intentional levels: (1) distal or future-oriented, (2) proximal, and (3) motor. Different cues to agency arise through the interaction of these different levels. Particularly interesting for the present purpose is the proximal level. Proximal intentions specify which perceptual events will occur as a consequence of an action, and it is likely that this level of detail is central for experiencing agency. However, the level of proximal intentions is subject to influences of higher-level plans and expectations and, at the same time, is affected by sensorimotor cues that are generated while actions are executed.

Pacherie's framework provides a highly useful integration of earlier theories that had focused on particular aspects of agency and brings out the fact that agency is influenced by cognitive factors as well as perceptual and sensorimotor cues to agency (cf. Sebanz & Lackner, 2007). The illusionary feelings of control that Wegner and co-workers are most interested in can be understood as strong cognitive influences arising on the distal level that overrule valid perceptual and sensorimotor cues to agency.

However, the framework also leaves ample room for theories that specify in detail how performance-related perceptual and sensorimotor cues affect one's feeling of causing events. (This can be considered as a special form of metacognitive judgment, see Metcalfe & Greene, 2007.) In accordance with Jeannerod and colleagues' notion of a who-system (Georgieff & Jeannerod, 1998; Jeannerod, 1999, 2003) perceptual cues can be linked to the proximal level. In accordance with notions that the monitoring of action execution is crucial for agency (Frith, 1992), sensorimotor cues can be understood as being derived from efferent signals (Tsakiris, Prabhu, & Haggard, 2006) or from discrepancies between the predicted and the actual sensory consequences of movements (Frith, Blakemore, & Wolpert, 2000) that are passed on from the motor level to the proximal level. In the following, we will briefly review earlier work addressing perceptual and sensorimotor cues to agency.

How can sensorimotor cues affect agency? One possibility is that efferent signals sent to the motor system while implementing an intention provide signals about how and when particular body parts should move (Tsakiris & Haggard, 2005; Tsakiris, Haggard, Franck, Mainy, & Sirigu, 2005). In support of this assumption Tsakiris and colleagues demonstrated that self-recognition of one's own body movements crucially depends on efferent signals. Furthermore, intentional movements and the resulting sensory events are perceived as being closer in time than passive movements and the events resulting from them (intentional binding, e.g., Haggard, Clark, & Kalogeras, 2002; Tsakiris & Haggard, 2003).

Another mechanism that has been proposed is that internal forward models derive a prediction of the sensory consequences of action from an efference copy that is issued with the actual motor command (Frith et al., 2000). Discrepancies between predictions and the actual incoming sensory input (re-efference) provide cues to external influences. Indirect evidence for this assumption comes from studies demonstrating that such discrepancies affect tactile sensations (Blakemore, Wolpert, & Frith, 1998, 2000) and visual perception of one's own actions (Leube et al., 2003). Direct evidence comes from studies that have explicitly demonstrated that agency is gradually reduced as these discrepancies increase due to spatial deviations or temporal delays (Fournieret & Jeannerod, 1998; Knoblich & Kircher, 2004; Leube et al., 2003; Sato & Yasuda, 2005; van den Bos & Jeannerod, 2002).

However, some of these results are open to alternative explanations that do not require sensorimotor cues. This is so because it is very hard to disentangle the action producing a perceptual event from the perceptual event itself, especially if the perceptual event is auditory or visual. At the same time it is well known that people tend to be unaware of their movements (e.g., Fournieret & Jeannerod, 1998) and that the phenomenology of action is very thin (cf. Vignemont, Tsakiris, & Haggard, 2006). So it could be that purely perceptual cues dominate agency. This would underline the central role of Pacherie's (2008) proximal level because it is the level that specifies which perceptual events are expected as a consequence of an intentional action.

Note that keeping apart perceptual events resulting from one's own and others' actions is not trivial on this level because they are not qualitatively different (Jeannerod, 1999; Knoblich & Flach, 2003). For instance, if the light goes on shortly after one has clicked the computer mouse one may experience the strange feeling that the mouse click turned on the light, although, in fact, somebody else pressed the light switch. This illustrates that perceptual events are often inherently ambiguous with respect to the actor who caused them.

Even if one postulates a who-system that attributes perceptual events to self and other (Georgieff & Jeannerod, 1998; Jeannerod, 1999, 2003; Vignemont & Fournieret, 2004; see also Ruby & Decety, 2001), misattributions of perceptual events to self can easily happen when another person is producing perceptual events one is expecting to happen as a consequence of one's own actions or when another person is matching the perceptual events she is producing to the perceptual events one is producing (remember the killer's footsteps at the beginning of the article).

Previous research has also addressed schizophrenic patients' pathological experiences related to agency such as auditory hallucinations, delusions of control, and thought disorders (Coltheart, 2007; Daprati et al., 1997; Frith, 1992; Frith et al., 2000; Jeannerod, 2009). It has been suggested that such experiences are due to a disturbed timing of efferent motor commands (Haggard, Martin, Taylor-Clarke, Jeannerod, & Franck, 2003), or faulty predictions generated by internal models in the motor system (e.g., Frith et al., 2000). In support of the latter assumption,

Lindner, Thier, Kircher, Haarmeier, and Leube (2005) have recently demonstrated that schizophrenic patients with delusions of control have problems with predicting the visual consequences of their own eye movements.

Other studies using manual movements suggest that perceptual (Fournier et al., 2002) and cognitive mechanisms (Knoblich, Stottmeister, & Kircher, 2004) are disturbed in patients who suffer from schizophrenia. For instance, Fournier and colleagues (2002) showed that schizophrenic patients were impaired in explicitly judging whether they were in control of perceptual events but were not impaired in automatically compensating for sensorimotor transformations between their movements and the resulting perceptual events (cf. Jeannerod, 2009).

1.2. The present study

As can be seen from the overview of previous work, there is evidence that cognitive expectations, perceptual cues, and sensorimotor cues can affect the sense of agency for the visible outcomes of actions. Each of these factors can contribute to pathological experiences of agency in schizophrenia (cf. Fletcher & Frith, 2009). Pacherie's (2008) conceptual framework for agency acknowledges that our sense of controlling our actions is influenced by different sources that are processed on different intentional levels. However, this framework leaves many specific questions open. For instance, the relative weight of evidence on each of the three levels with regard to experiencing agency is not specified. This raises the question to what extent future-oriented intentions or cognitive expectations can overrule perceptual and sensorimotor evidence, as suggested by Wegner (2002).

A second question is whether perceptual cues that are processed on the proximal level are more important than sensorimotor cues processed on the motor level. An important role of the motor level is suggested by findings demonstrating that people start compensating for sensorimotor transformations long before they are able to consciously detect the change (e.g., Fournier & Jeannerod, 1998; Knoblich & Kircher, 2004). Are there factors that can shift the balance between perceptual and sensorimotor cues? For instance, does ample expertise with a task enhance the influence of sensorimotor cues? This is suggested by the observation that many forms of expertise such as playing football or playing a musical instrument involve acquiring very particular sensorimotor mappings. As a consequence, experts do not only seem to be able to exert an amazing amount of control over actions in their domain of expertise but they also seem to feel more in control of these actions than novices.

A third related question is whether the balance between perceptual and sensorimotor cues is the same in the visual and auditory modalities. Does the exquisite timing that auditory perception provides imply a more important role for the perceptual level than in the visual domain? Does audition provide agency cues (e.g., related to temporal variability) that the visual system does not provide?

To answer these questions, we developed a paradigm that allows us to disentangle perceptual and sensorimotor cues to agency, and, to some extent, the influence of cog-

nitive expectations. Furthermore, we addressed agency in the auditory domain rather than in the visual domain, which had been addressed in most of the previous research (for exceptions, see Repp & Knoblich, 2004; Sato & Yasuda, 2005). Finally, we investigated whether skill level influences the relative contributions of perceptual and sensorimotor cues to agency. We achieved this by comparing the pattern of results obtained in a general sample of student participants with results for highly skilled musicians obtained in a previous study, which used a paradigm similar to that of Experiment 1 (Repp & Knoblich, 2007).

The experimental paradigm was inspired by earlier research on sensorimotor synchronization (Flach, 2005; Frassiné & Voillaume, 1971). The task consists in tapping with one's finger as regularly as possible while hearing tones at the same rate. The tones can either be triggered by the finger taps (i.e., controlled by the self, or *self-control*) or they can be controlled by the computer (i.e., controlled by an external agent, *external control*), in which case participants synchronize their taps with the tones. At some unpredictable point in a trial, one mode of control changes to the other mode.

There are four conditions: two active ones (changing from external-to-self or from self-to-external) and two passive ones (changing from external-to-self or from self-to-external). The active conditions are illustrated schematically in Fig. 1. In the active external-to-self condition, participants start out synchronizing their taps with a computer-controlled isochronous sequence of tones, but at some unpredictable moment control switches from external-to-self. They are asked to report when they detect that they are in control of the tones. Conversely, in the active self-to-external condition, participants start out in control of the tones, but at some point the computer takes over; they are asked to report when they detect that they have lost control over the tones. These two conditions are called *active* because tapping is involved. In the two *passive* conditions, participants sit still and listen to recordings of the tone sequences they heard in the active conditions. Again, they are asked to detect when control switched from external-to-self, or from self-to-external.

There are two principal cues to agency in this task (see Fig. 1). Asynchronies between taps and tones provide *sensorimotor* cues to agency because they involve a comparison between the time when an action occurs (as judged on the basis of tactile, kinesthetic, and visual information) and the time when its potential auditory consequence occurs. Detectable sensorimotor discrepancies can only arise during external control in the active condition, where asynchronies vary from tap to tap due to human movement variability. No detectable sensorimotor discrepancies exist during self-control, where asynchronies are small and constant (reflecting only an electronic processing delay; see Methods). A switch from external control to self-control thus implies a switch from presence to absence of sensorimotor discrepancies. Conversely, a switch from self-control to external control implies a switch from absence to presence of sensorimotor discrepancies.

Sensorimotor cues are absent in the passive condition because people are not tapping. Only *perceptual* cues re-

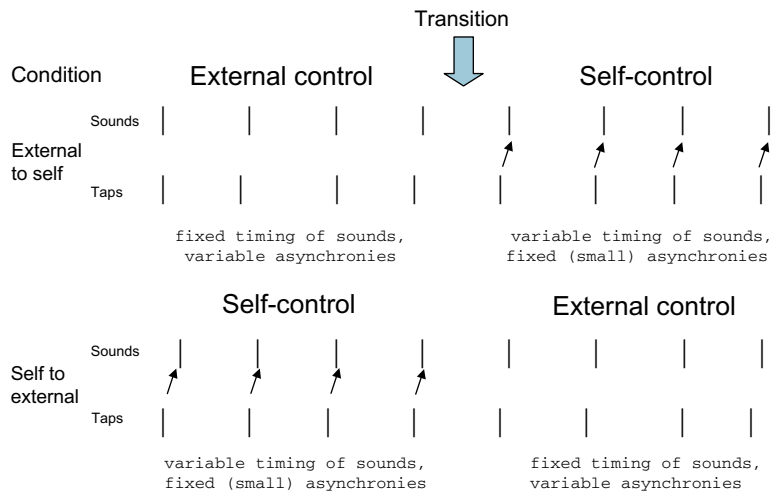


Fig. 1. Schematic illustration of the external-to-self and self-to-external conditions in Experiment 1. Eight tap-sound event pairs are shown for each condition, with the transition in control mode occurring after the fourth pair. The small fixed asynchronies during self-control are due to an electronic processing delay.

main, provided mainly by the temporal variability of the intervals between sounds. These intervals are variable during self-control because of the imperfection of human timing, but not during external control because the computer has been programmed to produce perfectly isochronous sequences. A switch from external control to self-control implies a switch from absence to presence of timing variability in the sound sequence, whereas a switch from self-control to external control implies a switch from presence to absence of timing variability. Of course, these perceptual cues are also present in the active condition.

By comparing detection performance in the active and passive conditions, we can determine to what extent sensorimotor cues are used to infer agency. In a previous study on music experts (Repp & Knoblich, 2007) we indeed found detection of a change in control mode to be easier in the active than in the passive conditions. In addition, while performance in the active external-to-self and self-to-external conditions was similar, the passive self-to-external condition was much more difficult than the passive external-to-self condition. This suggests that a transition from presence to absence of temporal variability in a sound sequence is harder to detect than the reverse transition, at least for highly trained musicians.

The efficacy of sensorimotor and perceptual cues was further confirmed in response-locked analyses of the data that showed that correct responses (hits) were immediately preceded by increases in asynchrony in the active self-to-external condition (note that responses usually occurred only after a number of taps in the new control mode had been made) and by increases in temporal variability in both active and passive external-to-self conditions. Thus the study effectively separated sensorimotor and perceptual contributions to inferring agency from sound. A further interesting result was that false alarm responses were more frequent during external control than during self-control. This may have been due to a general cognitive tendency to feel in control of events.

In the present study we investigated how sensorimotor cues, perceptual cues, and cognitive expectations contribute to the sense of agency in the general population of university students. Experiment 1 used our original paradigm, as described above, and thus was expected to reveal to what extent our earlier findings depended on musical expertise. Experiments 2 and 3 used a simplified paradigm to generalize the results, facilitate the task, and arrive at a test format that potentially could be used to examine the feeling of agency in special populations, such as schizophrenics and neuropsychological patients.

2. Experiment 1

Experiment 1 examined how people infer agency from sound. We expected that participants would use perceptual as well as sensorimotor cues to infer self-control or external control. Thus we predicted that detection rates for changes in control mode would be higher in the active conditions than in the passive conditions. Furthermore, the detection rates in the passive conditions should be above chance if participants make use of perceptual cues. According to the results of our earlier study with music experts (Repp & Knoblich, 2007), we further expected to find that detection scores would be lower in the passive self-to-external condition than in the passive external-to-self condition because the change from presence to absence of timing variability is particularly difficult to detect.

Earlier studies in the visual domain have demonstrated that for the average person perceptual cues can dominate over sensorimotor cues in the judgment of agency (e.g., Fournieret & Jeannerod, 1998; Knoblich & Kircher, 2004). If the same is true in the auditory domain, the difference between the active and passive conditions might be relatively small in the present study. In addition, poorer overall performance might be predicted for participants who are not expert musicians. Less extensive experience with rhythmic sound production and temporal discrimination

may be reflected in reduced sensitivity to temporal cues. However, this disadvantage might be partially offset by generation of larger cues. All cues result from variability in rhythmic movement timing, which is likely to be greater in people who are not musically trained. So if our participants generate larger cues for themselves, their overall performance might be similar to that of musicians.

Finally, we expected that participants might show a general tendency to attribute perceptual events to their own actions, particularly in the active conditions. Such effects have sometimes been observed in earlier studies on agency (e.g., Daprati et al., 1997). There were indications of such a cognitive bias towards self-attribution in our earlier study, and it is possible that musical non-experts are subject to stronger cognitive biases than expert musicians.

2.1. Methods

2.1.1. Participants

Twelve unselected Rutgers University undergraduates participated for course credit. We refer to them as musical non-experts because they clearly did not have the extensive musical training of our earlier group (Repp & Knoblich, 2007), though it is possible that some of them had some musical training. (Their level of training was not assessed.) The data of two participants were excluded because they did not follow instructions or because of technical problems.

2.1.2. Materials and equipment

Participants heard a snare drum sound, representing the first 100 ms of a digital sound file (“drumLoop.aif”) that accompanied the software used for running the experiment (MAX/MSP Version 4.5). This is one difference from our earlier study, in which a high-pitched digital piano sound was used instead. The MAX/MSP programs ran on a Macintosh G4 computer, and participants listened over Sennheiser HD250 linear II earphones.

2.1.3. Design and procedure

Participants came for two 1 h sessions on separate days, devoted to the external-to-self and self-to-external conditions, respectively. The order of the two conditions was counterbalanced across participants. Within each session, the active condition was followed (necessarily) by the passive condition. Because of time constraints, the number of trials was substantially smaller than in our earlier study. Participants received three blocks of 15 trials (i.e., a total of 45 trials) in each condition. The trials in each block resulted from the combination of three baseline inter-onset interval (IOI) durations (500, 600, and 700 ms) and five control mode transition locations (after the 8th, 12th, 16th, or 20th sound, or no transition). The different baseline IOI durations were included only for the sake of variety; they were not expected to interact with the results of interest (cf. Repp & Knoblich, 2007) and were ignored in the analysis of results.

External-to-self sequences consisted of 30 sounds. They started at one of the three baseline IOIs under external control and changed to self-control at one of the transition points. During external control, the timing of the sounds

was perfectly regular. During self-control, each sound was triggered by a tap but lagged behind the tap by approximately 42 ms (± 1 ms). The delay reflects electronic input and output processing, as no delay was specified in the MAX/MSP program. The delay was revealed by later acoustic measurements of the intervals between the impact sounds (thuds) of the taps and the onsets of the triggered sounds. It should be noted that, when synchronizing taps with a computer-controlled sequence, participants typically tap ahead of the sounds by several tens of milliseconds (see, e.g., Aschersleben, 2002). The lag during self-control thus helped reduce any abrupt change in the mean asynchrony at the transition point, which might have provided a local cue to the transition (cf. Flach, 2005). A lag of 42 ms is generally not detectable as such (Aschersleben & Prinz, 1997; Wing, 1977).¹

Self-to-external sequences started with three computer-controlled sounds at one of the three baseline IOIs, which served to induce a particular tempo of self-paced tapping. These induction sounds (not counted) were followed by 30 sounds that were first under self-control and then changed to external control at one of the transition points, or not at all. To avoid an abrupt change in mean IOI (i.e., a tempo discontinuity) at the self-to-external transition, the computer adopted the final inter-tap interval (ITI) before the transition as the constant IOI between sounds after the transition (external control).

Each active condition was preceded by five practice trials. In the passive conditions, the IOIs (specified during external control) or ITIs (registered during self-control) of the active conditions were used to regenerate nearly identical sound sequences. The interval between the last externally controlled tone and the first self-controlled tone was shortened somewhat in the passive condition because the 42 ms processing delay was not taken into account by the program.

The procedure was identical with that in Repp and Knoblich (2007). Participants sat in front of the computer screen on which a number (1–30) generated by a counter was displayed in a large font. The counter was incremented whenever a sound occurred. Participants started a trial by pressing the space bar and then closed their eyes. When they thought they had detected a transition, they opened their eyes briefly, glimpsed the number displayed on the screen, and closed their eyes again. At the end of the trial, they entered the remembered number on the keyboard. If no transition was detected, the response “30” (shown on the screen at the end of the trial) was to be given.

The nature of the transition to listen for was explained carefully at the beginning of each session. Participants were told that the transition could occur anywhere in a sequence or not at all. Two points were stressed especially: First, it was not permitted to revise any response on the ba-

¹ The lag was shorter (about 15 ms) in our previous study with musicians (Repp & Knoblich, 2007), due to the use of a digital piano for on-line sound generation. We did not mention the lag in our published article because at the time we were unaware of it. Because the tendency to tap ahead of sounds during synchronization is generally smaller in expert musicians than in non-experts (see, e.g., Repp & Doggett, 2007), the shorter delay was appropriate (if not exactly planned) in our experiment with musicians.

sis of later counterevidence; whatever number was seen first on the screen was to be the response, even if it turned out to be a false alarm. Second, participants were urged to tap as regularly as possible in the active conditions. Tapping always was to start with the third tone heard in a sequence. There were short breaks between blocks.

2.2. Results

The tasks were vulnerable to irregular participant behavior, and resulting technical problems led to the occasional loss of individual trials or even a whole block of trials. This affected the passive conditions in particular, because anomalous or missing responses in the active conditions could wreak havoc with the playback of the recorded sequences. All in all, 124 trials (6.7%) were lost for analysis.

From the usable trials in each of the four conditions for each individual participant, cumulative proportions of hits and false alarms were calculated as a function of sequence position. A numerical response was considered a hit if it referred to a sequence position that followed a transition; it was considered a false alarm if it referred to a position that preceded a transition, or to any position in a no transition catch trial. For each of the four transition locations, a separate cumulative function of hit proportions was obtained, starting at the transition point. A single cumulative function of false alarm proportions across all positions was obtained by combining false alarm responses from catch trials with those from transition trials.

From these data, four d' functions, one for each transition location, were obtained for each condition by calculating the difference between the z -transformed cumulative hit and false alarm proportions in each sequence position. These four d' functions were then re-expressed as functions of the serial distance from the transition and were averaged. The largest serial distance to which all transition locations contributed was 9; longer distances, whose data came from the earlier transition locations, provided little additional information and were not considered further. Finally, the individual d' functions for each condition were averaged across participants for graphic display. These average d' functions, one for each experimental condition, are shown in Fig. 2A.

It is clear that d' increased with distance from a transition in three of the four conditions. In the passive self-to-external condition, however, performance remained close to chance throughout. A repeated-measures ANOVA with the variables of transition type (external-to-self vs. self-to-external) and action condition (active vs. passive) was conducted on the individual d' values at a serial distance of 9. (The obviously significant main effects and interactions involving serial distance as a variable were of little interest.) The main effect of transition type was significant, $F(1,9) = 18.40$, $p < .002$, reflecting better detection performance for external-to-self than for self-to-external transitions. Unexpectedly, the main effect of action condition was not significant, but the two-way interaction was significant, $F(1,9) = 15.91$, $p < .003$: While performance in the active self-to-external condition was better than in the passive self-to-external condition, as predicted, perfor-

mance in the active external-to-self condition was actually worse than in the passive external-to-self condition.

As a measure of bias, we analyzed the z -transformed false alarm proportions. Although this measure is not independent of d' , other measures of response bias (c or β) would make use of hit proportions and thus would depend on transition location, which is not desirable in the present context. Fig. 2B shows bias as a function of sequence position in the four experimental conditions. Bias increased steadily within a sequence, which reflects the increasing expectation that a transition will occur and/or the increasing probability of perceiving a misleading cue.² It should be noted that a tendency to attribute control to the self results in a positive bias in external-to-self conditions (reporting self-control during external control) but in a negative bias in self-to-external conditions (not reporting external control during self-control). It can be seen that the self-attribution tendency was relatively greater in the active than in the passive conditions; that is, bias was larger in the active than in the passive external-to-self condition, and smaller in the active than in the passive self-to-external condition. A 2×2 repeated-measures ANOVA on the terminal bias values (sequence position 29) revealed no significant main effects but a significant two-way interaction, $F(1,9) = 8.59$, $p < .02$, which actually represents a main effect of action condition on the tendency to attribute agency to the self.

2.3. Discussion

The present results demonstrate that participants who are not highly trained musicians are able to use perceptual and sensorimotor timing cues to infer agency from sound, even though (as in our earlier study with musicians) these cues were not explicitly pointed out. The participants were quite able to use perceptual cues to agency when a transition occurred from external control to self-control in the passive condition. However, they performed at chance when they were asked to detect a switch from self-control to external control in the passive condition. Thus, while they were able to detect the presence of timing variability following its absence, they seemed unable to detect the absence of timing variability following its presence. The musicians in Repp and Knoblich (2007) had shown a similar difference between these conditions, but had performed at above chance levels.

The difficulty of detecting the disappearance of variability can be explained in at least two ways: First, undetectably low variability can occur by chance during self-control and thus is not a reliable cue to external control, whereas variability cannot arise spontaneously during external control (unless it is illusory) and therefore is a strong indicator of self-control. Second, exposure to temporal variability in a sound sequence decreases a participant's perceptual sensitivity to subsequent temporal variation (Large & Jones, 1999; Repp, 2002), probably by increasing the variability of an internal timekeeper or oscillator. This internal noise,

² The brief downward turn of the functions after position 20 is an artifact of combining false alarms from catch trials and transition trials. For unknown reasons, transition trials showed a higher cumulative false-alarm rate at this point than did catch trials.

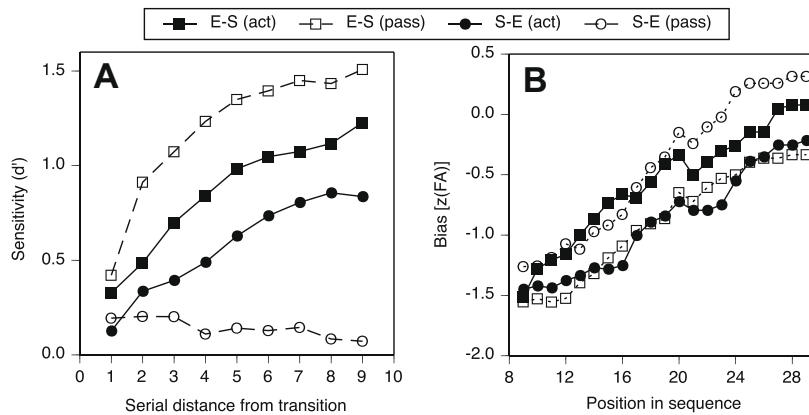


Fig. 2. (A) Sensitivity (d') as a function of serial distance from a transition in the four conditions of Experiment 1. (B) Bias (the z -transformed false alarm rate) as a function of sequence position in the same four conditions. E = external control; S = self-control; act = active; pass = passive.

too, makes it difficult to detect the absence of variability following self-control and may even lead to illusory perception of variability.

Clearly, sensorimotor cues were also used to infer agency from sound. This is apparent from the result that d' scores in the active self-to-external condition were higher than in the passive self-to-external condition. Thus the participants took the sudden appearance of asynchronies as a cue that they had lost control over the sounds. However, the present participants were insensitive to the sensorimotor cues that the active external-to-self condition provided, that is, the sudden disappearance of asynchronies. In fact, performance in the active external-to-self condition tended to be worse than in the passive external-to-self condition. This suggests that in the active condition it was harder for the participants to make use of the perceptual cues that were also present in the passive external-to-self conditions (the sudden emergence of temporal variability in the sound sequence). Performing the tapping task thus may have interfered with the processing of perceptual cues to agency.

The reasons for participants' difficulty in detecting the absence of asynchronies may be the same as we gave above to explain the difficulty in detecting absence of timing variability. Asynchronies, too, may become undetectably small during external control simply by chance, whereas they cannot arise spontaneously during self-control. This creates an asymmetry in the information conveyed by absence versus presence of asynchronies. Also, it is possible that exposure to asynchronies during external control increases internal variability and therefore makes it difficult to discriminate between presence and absence of asynchronies.

Differences in the pattern of d' scores between the present group of participants and the highly trained musicians in our earlier study (Repp & Knoblich, 2007) thus seem to be mainly due to the fact that the present group was insensitive to transitions from the presence to the absence of perceptual or sensorimotor variability. In addition, tapping seemed to interfere with detection of perceptual cues in the present group. Furthermore, musicians showed greater sensitivity to agency cues overall, performing clearly better

than the present group in three out of four conditions; only in the passive external-to-self condition were the d' scores similar. We had expected that larger variability of tapping would compensate for poorer perceptual acuity in the present group, but this compensation may have been only partial. However, it is difficult to compare overall performance between the studies because other factors may have affected performance as well. In particular, the musicians received many more trials and thus benefited from practice in the tasks; they had previous experience with synchronization tasks; and they were also highly motivated, whereas the present participants merely fulfilled a course requirement. Furthermore, the drum sound, because of its lower frequency content, was probably a more effective masker of the thuds made by the taps on the tapping pad than was the high-pitched piano tone used in our study with musicians. Thus, the present participants may have had to rely exclusively on cross-modal information in perceiving asynchronies between taps and sounds, whereas the musicians may have had the benefit of some asynchrony information within the auditory modality.

Our analysis of bias showed the predicted tendency to attribute control of events to the self in the active conditions, relative to the passive conditions. This is consistent with the hypothesis that people have a general bias towards feeling in control when they act. However, the evidence is not conclusive because our measure of bias is not independent of sensitivity and thus could also reflect differences in the effectiveness of sensorimotor cues. Feelings of self-control during active synchronization may have arisen because asynchronies were hard to detect, not because of a cognitive bias. Reports of external control during active self-paced tapping may have been relatively infrequent because the absence of perceived asynchronies reinforced the impression of self-control.

The main difference between the present bias results and the bias results for the group of musicians (Repp & Knoblich, 2007) is that the musicians showed much smaller false alarm rates in the active conditions; bias in the passive conditions was quite similar for the two groups. The musicians did not show any greater tendency to report self-control during synchronization; however, they were

much less likely to report external control during self-paced tapping than were the present participants. These differences in false alarm rates obviously contribute to the differences in d' scores between the two groups.

3. Experiment 2

In Experiment 2 our aim was to simplify the paradigm in order to reduce the difficulty of the task, and to avoid some of the complexities of analysis and interpretation evident in Experiment 1 and in our previous study. Simplification also served the goal of developing a version of our agency task that could potentially be administered to neurological and psychiatric patients.

The task we used in Experiment 1 has several disadvantages, especially from a practical viewpoint. First, the ability to detect a transition between control modes is not a very direct measure of agency. In the present experiment, we simply asked participants to decide after each trial whether they had been in control or not. Second, we had to trust participants to follow the instruction always to tap as regularly as possible, even though this made the task difficult for them. If some participants had decided not to follow instructions and tapped irregularly, this would have greatly facilitated their agency judgments. Even though this problem could not be eliminated entirely, being inherent in agency paradigms, we conducted a more rigorous check on such undesirable strategies in Experiments 2 and 3. Third, our measure of bias in Experiment 1 was not independent of d' , which resulted in inconclusive findings regarding the influence of cognitive expectations on inferring agency from sound. In the new simplified paradigm we were able to use a bias measure (c) that is independent of d' . Finally, the sequences we used in Experiment 1 were longer than necessary: Most of the relevant information came from the early sequence positions following a change in control mode. Therefore we decided to shorten the sequences radically.

In order to implement these changes, we employed a sequence reproduction task: A short model sequence of isochronous sounds was presented, and the participant's task was to listen to it and then to reproduce it as precisely as possible by tapping. The taps either triggered sounds (self-control) or were accompanied by a computer-controlled reproduction of the sequence that was triggered by the first tap (external control, except for the first sound). We still had active and passive conditions. In the active condition, perceptual and sensorimotor cues were available and participants reported whether the sounds heard during reproduction were self-controlled or externally controlled. In the passive condition, participants listened to a playback of the model sequences followed by their recorded reproductions and judged whether the reproductions had been self-controlled or externally controlled.

Our predictions were the same as in Experiment 1. We expected better discrimination of external control and self-control (i.e., higher d' scores) in the active condition than in the passive condition due to the presence of sensorimotor cues to agency (presence vs. absence of asynchronies) in the former condition, in addition to perceptual cues (ab-

sence vs. presence of timing variability in the sound sequence) that were available in both conditions. Moreover, we expected d' scores to be above chance in the perceptual condition, indicating that perceptual cues can be used to infer agency from sound. Finally, we expected to find a general cognitive tendency to attribute control to the self in the active condition, compared to the passive condition.

3.1. Methods

3.1.1. Participants

Sixteen unselected Rutgers University undergraduates participated for course credit.

3.1.2. Materials and equipment

The drum sounds and the equipment were the same as in Experiment 1. The only difference was that participants, except for the first few, listened to the sounds not over earphones but over the computer's built-in loudspeaker, located in the processor standing to the right of the monitor the participants were facing. This change was made to enable the experimenter to monitor participants' performance during the experiment and intervene if necessary (e.g., when taps were omitted or tapping was clearly irregular).

3.1.3. Design and procedure

A model sequence consisted of 3, 4, 5, or 6 drum sounds played with fixed IOIs of 400, 500, or 600 ms. Both variables were included only to provide variety and were ignored in subsequent analyses. Their combination yielded 12 sequence types that occurred in both external control and self-control trials, for a total of 24 randomly ordered trials per block. Three such blocks (a total of 72 trials) were presented in the active condition, preceded by two practice blocks. To aid participants in producing the right number of tones, a digit indicating the number of tones was shown on the computer monitor during each trial.

The first practice block contained 12 self-control trials in which the drum sounds during reproduction were always contingent on the participant's taps. Depression of the space bar on the computer keyboard started a trial. Participants were instructed to listen to the model sequence and, after a brief pause, to reproduce the model sequence as accurately as possible by tapping. The MAX program provided verbal feedback after each trial on the computer screen: If the mean ITI of the reproduction was within $\pm 10\%$ of the model IOI, the feedback was "Good!"; if not, the feedback was either "Too slow!" or "Too fast!".

The second practice block contained 12 external control trials in which the reproduction of the model sequence was controlled by the computer. Participants were instructed to listen to the model sequence and, after a brief pause, to reproduce the model sequence in synchrony with the computer-controlled sounds. If the maximal absolute asynchrony between taps and drum sounds was less than 20% of the model IOI, the feedback was "Good!"; otherwise, the feedback was "Not accurate enough!".

During the three test blocks, self-control and external control trials occurred in random order, and no feedback was given. After each trial, the participant had to respond

to the question “Who controlled the second series of sounds?” by clicking on one of six numbered and labeled response buttons on the computer screen, effectively a 6-point rating scale. The labels ranged from “I did (definitely)” to “The computer did (definitely)”, with intermediate steps being labeled “I did (probably)”, “I did (perhaps)”, “The computer did (perhaps)”, and “The computer did (probably)”. The response started the next trial after a delay of 3 s. There were short breaks between blocks during which the data were saved and the next block was initiated by the experimenter.

In the subsequent passive condition, the IOIs or ITIs recorded during the active condition were used to regenerate the sound sequences (both models and reproductions). At the end of each trial, participants responded to the same question as in the active condition, using the same response buttons. An additional button was provided to give participants an opportunity to disqualify trials in which the reproduction sounded anomalous (too few or too many sounds, unusually long intervals, etc.), but it was rarely used.

3.2. Results

Before analyzing the results, each participant’s data were screened to eliminate all trials in which there were missing or extra taps, all self-control trials in which the mean ITI of the reproduction deviated by more than ± 100 ms from the model IOI or in which the standard deviation of the reproduced intervals exceeded 100 ms, and all external control trials in which the mean asynchrony exceeded ± 100 ms or in which the standard deviation of the asynchronies exceeded 100 ms. (These tolerance limits were arbitrary but relatively lenient.) For three participants, this left less than 70% of the trials, and their data were excluded entirely. For the remaining 13 participants, the percentage of trials analyzed ranged from 73% to 97%.

Few participants used the 6-point rating scale effectively; many of them used only the endpoints. Therefore, the responses were dichotomized by splitting the scale in the middle, between 3 and 4. In two cases, however (the active condition for one participant and the passive condition for another), the scale was split between 2 and 3 and between 5 and 6, respectively, in order to prevent all re-

sponses from forming a single category. From the resulting proportions of hits and false alarms in each condition for each participant, combined across all IOIs and numbers of tones in the sequence, d' and the bias index c (the mean of the z -transformed hit and false alarm proportions) were calculated.

Fig. 3A shows the mean d' values. Discrimination of self-control from external control was better in the active than in the passive condition, $F(1,12) = 7.24$, $p < .02$. The mean bias indices are shown in Fig. 3B. In the active condition, there was a positive bias, indicating a tendency to report self-control. The bias in the passive condition was significantly smaller, $F(1,12) = 19.08$, $p < .001$, and not significantly different from zero.

The data were further analyzed to examine the cues on which participants presumably relied in making their responses. Fig. 4A and B show two positive perceptual cues to self-control: the mean absolute difference between the reproduction ITIs and the model IOI (i.e., the mean tempo difference between model and reproduction, Fig. 4A) and the mean standard deviation of the reproduction ITIs (i.e., temporal variability, Fig. 4B). These variables are shown separately for the active and passive conditions, and contingent on a response of self-control or external control within conditions. It is clear that both cues were more pronounced when participants responded “self” (“S”) rather than “external” (“E”), and this difference was larger in the passive than in the active condition. A 2×2 repeated-measures ANOVA on the data in Fig. 4A revealed a significant main effect of response, $F(1,12) = 14.58$, $p < .002$, and a significant interaction of response with condition, $F(1,12) = 8.83$, $p < .02$. Separate one-way ANOVAs on the two conditions showed the effect of response to be significant in the passive condition only, $F(1,12) = 27.69$, $p < .001$. A similar ANOVA on the data in Fig. 4B likewise showed a significant main effect of response, $F(1,12) = 6.14$, $p < .03$, but the interaction fell short of significance, $F(1,12) = 4.16$, $p < .07$. Nevertheless, separate ANOVAs again showed the effect of response to be significant only in the passive condition, $F(1,12) = 8.35$, $p < .02$.

Fig. 4C and D show two positive sensorimotor cues to external control: the mean absolute asynchrony (Fig. 4C) and the mean standard deviation of the asynchronies (Fig. 4D). In each case, these variables were larger when

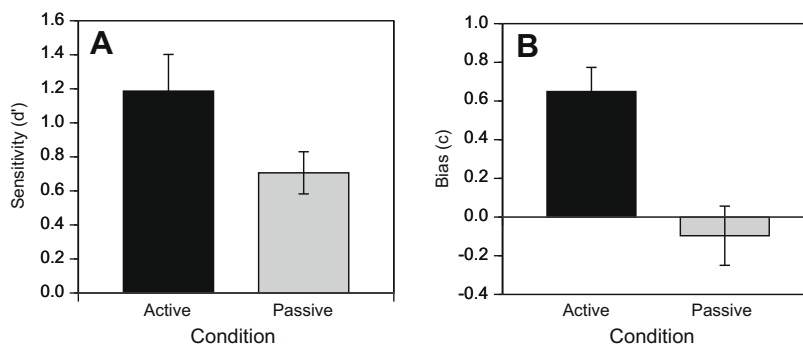


Fig. 3. (A) Mean sensitivity (d') in the two conditions of Experiment 2. (B) Mean bias (c) in the same two conditions. The error bars represent standard errors.

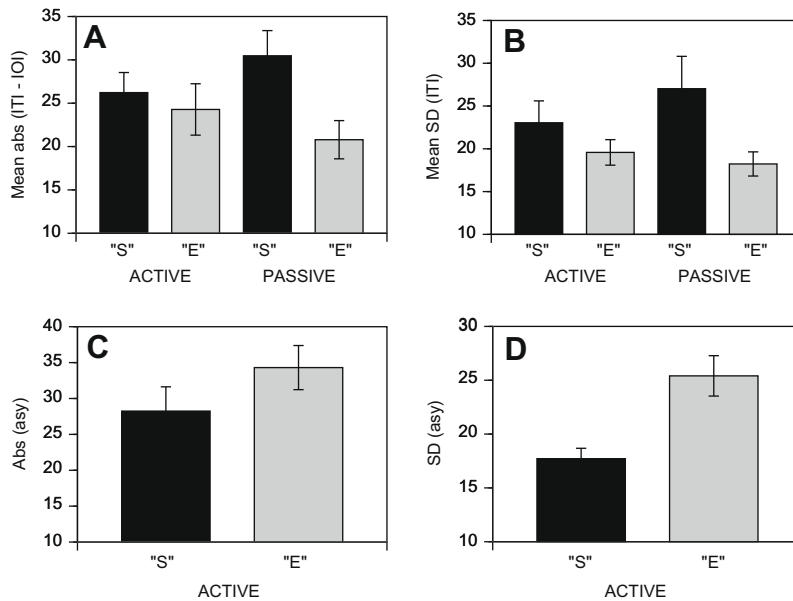


Fig. 4. Response-locked cue analysis in Experiment 2. Self-control: (A) Mean absolute difference between reproduction ITIs and model IOI and (B) mean standard deviation of ITIs, given "self" ("S") and "external" ("E") responses in the active and passive conditions. External control: (C) Mean absolute asynchrony and (D) mean standard deviation of asynchronies, given "self" ("S") and "external" ("E") responses in the active condition. The error bars represent standard errors.

the response was "external" than when it was "self". However, the difference was significant only for the data in Fig. 4D, $F(1,12) = 18.32$, $p < .001$.

3.3. Discussion

As expected, discrimination of self-control from external control was easier in the active than in the passive condition. This clearly reflects the exclusive availability of sensorimotor cues in the active condition. Of the two likely sensorimotor cues for external control that we considered in the response-contingent analysis, the standard deviation of the asynchronies, which reflects the temporal variability of the taps, appeared to be more salient than the mean absolute asynchrony, which reflects the overall tendency to tap before (or after) the computer-controlled sounds.

Perceptual cues to self-control, which were available in both active and passive conditions, were more effective in the passive than in the active condition. They were somewhat overshadowed by sensorimotor cues in the active condition, as indicated by the lack of a significant effect in the response-contingent analyses of perceptual cues. The two perceptual cues examined, the tempo difference between model and reproduced sequences and the temporal variability of the reproduction, seemed to be about equally salient.

In the active condition, there was a strong tendency to report self-control, which was absent in the passive condition. This is consistent with the prediction that people have a general cognitive bias towards feeling in control when an action is carried out. However, as in Experiment 1, there is still an alternative explanation in terms of an informational asymmetry of sensorimotor cues: During external control,

asynchronies often may have been too small to detect, which led to a false report of self-control. During self-control, however, participants would have had to hallucinate asynchronies in order to conclude falsely that they were not in control. Thus, to the extent that participants focused on asynchronies in making their decisions, the bias towards self-control follows naturally.

It could be argued that a complementary informational asymmetry of perceptual cues should have counteracted the self-control bias in the active condition. An inability to detect temporal variation in the sound pattern during self-control should have led to incorrect external control responses, whereas only a hallucinatory perception of variability could have led to incorrect self-control responses during external control. However, the response-contingent cue analyses suggested that perceptual cues were somewhat overshadowed by the sensorimotor cues in the active condition, and this could account for the overall self-control bias.

Furthermore, it may indeed be easier to hallucinate temporal variability in a regular sound sequence than to hallucinate asynchronies between coincident taps and tones. Actions and their potential consequences tend to be bound together temporally (Haggard et al., 2002), which also makes it hard to detect small asynchronies when they are present. This binding effect may also have contributed to the response bias observed in the active condition, and it is not itself a cognitive bias, although it can be influenced by cognitive factors (Engbert & Wohlshläger, 2007). In any case, it is possible that the response bias towards self-control observed in the active condition is not due to cognitive expectations but to the asymmetries in sensorimotor cues discussed above.

4. Experiment 3

The aim of our third experiment was to replicate the findings of Experiment 2 with an even more simplified sequence reproduction task. Furthermore, we attempted to further explore the role of asynchronies in creating the response bias favoring self-control that arose in the active condition of Experiment 2. At the same time, we wanted to facilitate the task for the participants, for d' scores were still rather low in Experiment 2. To that end, we introduced tempo differences between the model sequence and the computer-controlled reproduction during external control, while instructing participants to tap their reproduction at the model sequence tempo. We expected that larger tempo differences would create larger asynchronies. Accordingly, d' scores should increase in the active condition. Moreover, if the response bias observed in the active condition of Experiment 2 was related to a difficulty in detecting asynchronies it should become smaller as tempo differences become larger. If the response bias was due to cognitive expectations, however, tempo differences should not matter.

We expected that the introduction of tempo differences during external control would also affect participants' ability to infer agency from sound in the passive condition. Specifically, because tempo differences were ruled out as a valid perceptual cue to self-control, participants had to rely on temporal variability as the only perceptual cue, and this was expected to increase the difficulty of the passive condition compared to Experiment 2. Thus, the difference between the d' scores in the active and passive conditions was expected to be larger in Experiment 3 than in Experiment 2.

4.1. Methods

4.1.1. Participants

Fourteen unselected Rutgers University undergraduates participated for course credit.

4.1.2. Materials and equipment

The sounds and the equipment were the same as in Experiment 2.

4.1.3. Design and procedure

In this experiment the model sequence always consisted of four drum sounds, played with IOIs of 400, 500, or 600 ms. For each of these three tempo conditions, there were five reproduction tempi during external control: The reproduction IOI was either the same as the model IOI or it was different by -10% , -5% , 5% , or 10% . This design led to 15 trial types for external control but only three for self-control; the latter were repeated 5 times each. Each of the three test blocks thus contained 30 randomly ordered trials (a total of 90 trials). The test blocks were preceded by two practice blocks, for self-control and external control, respectively, each of which contained 15 trials.

During self-control practice, participants were asked to reproduce the model sequence as precisely as possible. During external control practice, they were given the same

instructions but in addition were told to *ignore* the sounds they heard. (In Experiment 2 they had been asked to synchronize their taps with the sounds.) They were informed that the computer would often play back the model sequence at a different tempo. In the test blocks, participants were likewise urged to reproduce the model sequence always at the original tempo. Relevant feedback was provided in all practice *and* test blocks (another difference from Experiment 2): The message "Good!" appeared when the mean ITI was within $\pm 10\%$ of the model sequence IOI; otherwise, the message was either "Too fast!" or "Too slow!"

The response was also simplified relative to Experiment 2. In response to the question "Who controlled the second series of sounds?" participants clicked one of two labeled buttons on the screen: "I did" or "The computer did".

As in Experiment 2, a passive condition followed the active condition. Participants were alerted to the fact that a change in tempo between the model and reproduced sequences in a trial was not a reliable cue to either self-control or external control, and that they should listen for the presence of temporal variability.

4.2. Results

Three participants performed at chance level in both the active and passive conditions; their data were excluded. The data of the remaining 11 participants were screened to eliminate all trials in which there were missing or extra taps, whose mean ITI deviated by more than 100 ms from the model sequence IOI, or in which the standard deviation of the ITIs exceeded 100 ms. This left 66% and 74% of the trials, respectively, for two participants (who were retained), and between 86% and 99% for the remaining 9 participants.

The main results are shown in Fig. 5. It is evident that mean sensitivity (d') was well above chance level in the active condition but basically at chance level in the passive condition (Fig. 5A). The difference between the two conditions was significant, $F(1,10) = 25.96$, $p < .001$. Surprisingly, performance in the active condition was not significantly better than in Experiment 2, although performance in the passive condition was worse (compare with Fig. 3A). A combined ANOVA of the two experiments showed a significant experiment \times condition interaction, $F(1,22) = 5.71$, $p < .03$, in addition to a significant main effect of condition, $F(1,22) = 33.08$, $p < .001$. Thus, the difference between the active and passive conditions was larger in Experiment 3, as predicted.

In contrast, the bias results (Fig. 5B) were quite similar to those of Experiment 2 (see Fig. 3B). Again, there was a significant self-attribution bias in the active but not in the passive condition. The difference between the two conditions was significant, $F(1,10) = 7.19$, $p < .03$.³

³ One participant gave only "self" responses in the passive condition, which resulted in a maximal bias value of 2.13 (response proportions of 0 or 1 were adjusted by adding or subtracting $1/2N$, respectively), far higher than any other participant's. Although this value was included in the statistical analysis (hence the low p -value), it was excluded from the mean value shown in Fig. 5B.

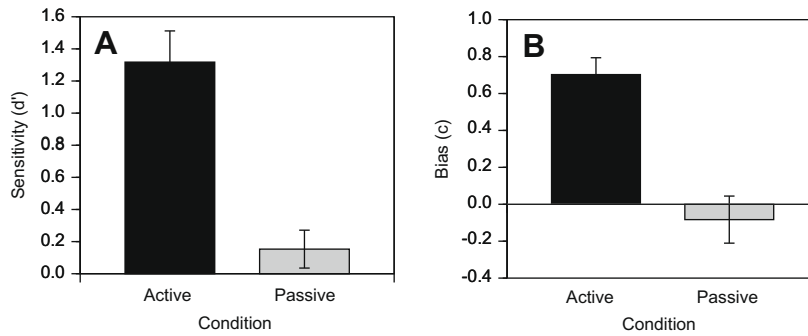


Fig. 5. (A) Mean sensitivity (d') in the two conditions of Experiment 3. (B) Mean bias (c) in the same two conditions. The error bars represent standard errors.

To determine whether the tempo changes in computer-controlled sequence reproductions affected performance, d' and c were computed for each tempo change condition separately in the active and passive conditions. Because individual data were sparse at this level of detail, the data of all participants were pooled before computing d' and c . In this analysis, a single hit proportion was obtained from self-control trials, and a separate false alarm proportion (self-control responses) was obtained from external control trials in each tempo difference condition. The d' and c scores thus reflect directly the variation in false alarm rates. Because of the pooling of data, no statistical analyses could be conducted on these results.

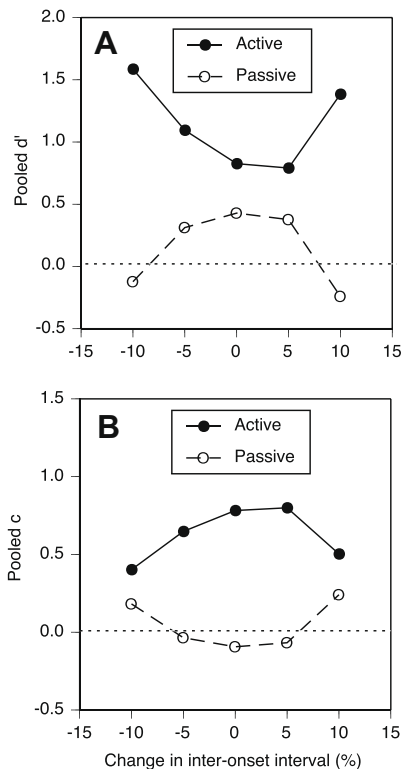


Fig. 6. (A) Pooled d' and (B) pooled c as a function of the tempo change between the model sequence and the computer-controlled reproduction in the active and passive conditions.

Fig. 6A shows d' as a function of tempo change. As was fully expected, discrimination of self-control and external control was easier in the active condition when there was a tempo change because this resulted in larger asynchronies. However, the larger tempo changes ($\pm 10\%$) also had an effect in the passive condition, where they pushed discrimination below chance level.

The bias coefficients, shown in Fig. 6B, follow the opposite pattern because they reflect the same false alarm rates. It can be seen that better discrimination performance in the active condition went along with a reduced tendency to report self-control when there was a large tempo change. Conversely, poorer discrimination performance in the passive condition went hand in hand with an increased tendency to report self-control when there was a large tempo change. Even though participants had been instructed to disregard tempo differences in the passive condition, they apparently still felt that such a change might have been due to human agency. The tempo change was also probably easier to detect than the presence of temporal variability, considering that performance in the passive condition was at chance overall.

We did not conduct any response-contingent analysis of cues, as we did in Experiment 2, because it was rather obvious that asynchronies were the salient cue in the active condition, whereas temporal variability was hardly detectable in the passive condition and therefore presumably also ineffective as a cue in the active condition.

4.3. Discussion

Experiment 3 replicated two main findings of the previous experiment: better discrimination between self-control and external control in the active condition than in the passive condition, and a self-attribution bias in the active condition only. The finding of very poor discrimination in the passive condition indicates that participants were unable to make use of perceptual cues to agency. Whereas tempo change cues had been invalidated by the design of the experiment, changes in timing variability simply seemed too hard to detect. This may explain why d' scores in the active condition of Experiment 3 were not higher than in the active condition of Experiment 2, despite the fact that larger asynchronies were created by the tempo changes in Experiment 3: The advantage of more salient

sensorimotor cues seems to have been offset by the lack of additional informative perceptual cues.

Despite participants' greater reliance on sensorimotor cues in the present experiment, the bias results were very similar to Experiment 2. Thus the bias in the active condition remained the same despite the larger asynchronies. This seems to favor the cognitive explanation that people have a general bias to feel in control of events when they are performing actions. However, the bias did decrease slightly as the asynchronies increased as a function of the tempo differences. This could be taken as support for the notion that the bias has its origin in the difficulty of perceiving asynchronies. Participants' known tendency to perceive actions and their consequences as close together in time (Haggard et al., 2002) may contribute to that difficulty. However, there was still a positive bias even when the tempo difference was large, suggesting that the self-attribution bias may reflect both cognitive and perceptual factors.

5. General discussion

How do people infer agency from sound? The three experiments we have reported suggest that there is no single sensorimotor or perceptual cue and no single thought that makes us sense or think that we have caused a sound that we hear. Rather, we seem to flexibly combine the information provided by sensorimotor and perceptual cues with cognitive expectations about the consequences of our actions.

All three experiments yielded evidence that the asynchronies between actions (taps) and their consequences (sounds) can provide a major sensorimotor cue to inferring agency from sound. In Experiment 1 participants could only detect a switch from self-control to external control when variable asynchronies were present during external control (active condition) and not when they were absent (passive condition). The results obtained with a new sequence reproduction task in Experiments 2 and 3 further underlined the important role of variable asynchronies for inferring agency. In both experiments, participants were better able to discriminate between self-control and external control when they were actively tapping than when they were passively listening. In fact, sensorimotor cues seemed to be the only available cues in Experiment 3 where discrimination was at chance level for passive listening.

However, participants' inability to discriminate between self-control and external control in the passive condition of Experiment 3 does not imply that people generally cannot use perceptual cues to agency. It is true that in Experiment 3 timing variability did not seem to be detected by participants, whereas tempo change cues had been invalidated by the design. In the passive listening condition of Experiment 2, however, tempo change could act as a valid cue, and participants were well able to distinguish between self-control and external control, implying that changes in tempo did provide a perceptual cue to agency.

Moreover, our analyses suggested that timing variability also acted as a perceptual cue in Experiment 2, which

suggests that the change in tempo from trial to trial in Experiment 3 interfered with the detection of timing variability within a given trial. Experiment 1 even suggested a preferential role of perceptual cues when control switched from external control to self-control: Participants were better able to detect the transition in the passive listening condition than in the active tapping condition. In our view, this unexpected result can be best explained if one assumes that tapping constituted a dual task for participants that actually kept participants from monitoring perceptual cues to agency. In any case, the relatively high d' scores in the passive condition of Experiment 2 illustrate that perceptual cues were definitely used to infer agency from sound.

The results are somewhat less clear with regard to the expected tendency to attribute control of perceptual events to self when acting. Although all three experiments provided evidence of such a tendency, its attribution to purely cognitive factors remains uncertain. Our experimental paradigms were developed to effectively dissociate sensorimotor and perceptual cues; at present, they cannot distinguish as reliably between cognitive and perceptual causes of bias. An interesting possibility is that intentional binding of sounds and actions (Haggard et al., 2002) contributes to the bias by making asynchronies harder to detect. Clearly, further studies are needed to disentangle the role of cognitive influences and different types of perceptual and sensorimotor cues to agency.

What are the wider implications of the present results? In addition to clearly separating the contributions of perceptual and sensorimotor cues to agency that were often confounded in previous studies, we have identified new cues to agency. These include (1) tempo differences between perceptual event sequences, (2) temporal variability of perceptual event sequences, and (3) temporal variability of asynchronies between movements and their auditory consequences. Although a previous study has shown that the latter two cues help expert musicians to establish agency (Repp & Knoblich, 2007), the present study is the first to demonstrate that musical expertise is not a precondition to make use of them. Identifying the new cues became possible because our new experimental setups allowed us to address agency for sequences of events rather than agency for discrete events that was addressed in previous studies.

The main theoretical implication of the above findings is that models of agency need to add mechanisms that accumulate evidence about sensorimotor and perceptual discrepancies across different actions and perceptual events. This has so far not been specified because it was assumed that discrepancies are computed between distinct actions and events (e.g., Frith et al., 2000; Jeannerod, 2009; Pacherie, 2008). An open issue that needs to be further explored is whether temporal variability can be used as a cue to agency in other perceptual modalities such as vision and touch.

A second novel aspect of the present results is that skill level and task difficulty can affect the relative weight that is given to different types of agency cues. In particular, the results suggest that actors rely more on perceptual cues if a task is difficult or unfamiliar. In the relatively difficult task

used in Experiment 1, music experts in our earlier study (Repp & Knoblich, 2007) were much more sensitive to sensorimotor cues to temporal variability than the present “ordinary” participants. However, in the less challenging task used in Experiment 2 ordinary participants also relied on sensorimotor cues, and even ignored perceptual cues in Experiment 3 when these cues became less informative.

An important theoretical implication of these findings is that, at the onset of acquiring a skill, agency could be derived from discrepancies on a perceptual planning level (distal level in Pacherie's, 2008, terms; cf. Jeannerod, 2009). However, if people perform particular difficult tasks over and over again, like pianists who learn to rhythmically press keys or drummers who learn to rhythmically hit drum pads, they will establish internal models that allow them to better control the fine-grained timing of their movements with respect to the perceptual events they produce. Thus these internal models provide additional sensorimotor cues to agency (cf. Frith et al., 2000). Of course, in other (easier) tasks everybody is a sort of expert because movement always involves timing. This could explain why ordinary participants were able to effectively use sensorimotor cues in Experiment 2 and 3.

The hypothesis that, at the onset of skill acquisition, agency is inferred on the basis of perceptual cues and only later inferred on the basis of sensorimotor cues, opens up interesting perspectives for future research. For instance, studies might be conducted in which people are trained to perform particular tasks that provide perceptual and sensorimotor cues to agency and afford implementation of new internal models. Will agency depend more on sensorimotor cues and less on perceptual cues, as people get better at the task? Similar experiments in children could be even more telling. Do kids who learn to bike feel control over their bikes, although their parents are needed in order to keep them in balance (implying that agency is inferred from perceptual cues)? If they switched back to riding a tricycle they have been using for a couple of years, would they then (also) rely on sensorimotor cues to infer agency?

Another noteworthy aspect of the results is that they provide further support for the assumption that cognitive expectations, beliefs, and thoughts can affect how much control people feel over actions and events (cf. Wegner, 2002) and that these influences seem to be relatively independent of task difficulty and expertise. The mere fact of performing actions in the active condition led people to consistently exhibit a bias towards reporting self-control that was not present in the passive condition. Although this bias may partly be caused by perceptual factors, it is unlikely that it is entirely of perceptual origin. An open question that still needs to be further explored is whether cognizing about the control one exerts over movements and their perceptual consequences is different from cognizing about causality in general. Linking agency research to related research on causal reasoning (e.g., Hagmayer, Sloman, Lagnado, & Waldmann, 2007) would likely help to make progress on this matter.

Finally, although we did not directly address delusions of agency in schizophrenia with the present experiments, the current findings have potential implications for future

research in this field. Although much progress has been made in understanding the neural bases of delusions (Fletcher & Frith, 2009), clinical studies could profit from tasks that allow one to separate the contributions of different levels of control and different agency cues in different modalities. In the best case, this would allow one to link particular forms of delusions to particular functional deficits (e.g., inability to process sensorimotor cues in the auditory domain).

To conclude, unlike many previous studies of agency, the present study focused on the role of audition rather than vision. We suggest that sounds may sometimes be more informative and important than visual cues for determining agency. One could argue that audition is the perfect modality to infer agency because of its high temporal accuracy and broad coverage of environmental events. If a person walks behind me, I simply do not see that person until I turn around. However, one can often infer that others are present from hearing their footsteps. Unfortunately, the sounds that make people feel in control of their footsteps also provide a potential cover for ruthless killers.

Acknowledgments

This research was supported by a Faculty grant from Rutgers University to Günther Knoblich and was conducted while both authors were affiliated with the Department of Psychology, Rutgers University, Newark, NJ. We thank Natalie Sebanz for many helpful comments.

References

- Aarts, H., Custers, R., & Wegner, D. M. (2005). On the inference of personal authorship: Enhancing experienced agency by priming effect information. *Consciousness and Cognition*, *14*, 439–458.
- Aschersleben, G. (2002). Temporal control of movements in sensorimotor synchronization. *Brain and Cognition*, *48*, 66–79.
- Aschersleben, G., & Prinz, W. (1997). Delayed auditory feedback in synchronization. *Journal of Motor Behavior*, *29*, 35–46.
- Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*, 635–640.
- Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2000). Why can't you tickle yourself? *NeuroReport*, *11*, R11–R16.
- Coltheart, M. (2007). Cognitive neuropsychiatry and delusional belief. *Quarterly Journal of Experimental Psychology*, *60*, 1041–1062.
- Daprati, E., Franck, N., Georgieff, N., Proust, J., Pacherie, E., Dalery, J., et al (1997). Looking for the agent: An investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition*, *65*, 71–86.
- Engbert, K., & Wohlschläger, A. (2007). Intentions and expectations in temporal binding. *Consciousness and Cognition*, *16*, 255–264.
- Flach, R. (2005). The transition from synchronization to continuation tapping. *Human Movement Science*, *24*, 465–483.
- Fletcher, P. C., & Frith, C. D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, *10*, 48–58.
- Fourneret, P., de Vignemont, F., Franck, N., Slachevsky, A., Dubois, B., & Jeannerod, M. (2002). Perception of self-generated action in schizophrenia. *Cognitive Neuropsychiatry*, *7*, 139–156.
- Fourneret, P., & Jeannerod, M. (1998). Limited conscious monitoring of motor performance in normal subjects. *Neuropsychologia*, *36*, 1133–1140.
- Fraisse, P., & Voillaume, C. (1971). Les repères du sujet dans la synchronisation et dans la pseudo-synchronisation. [The subject's references during synchronization and pseudo-synchronization.] *L'Année Psychologique*, *71*, 359–369.
- Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Hillsdale, NJ: Erlbaum.

- Frith, C. D., Blakemore, S.-J., & Wolpert, D. M. (2000). Abnormalities in the awareness and control of actions. *Philosophical Transactions of the Royal Society of London: Series B*, 355, 1771–1788.
- Georgieff, N., & Jeannerod, M. (1998). Beyond consciousness of external reality: A “Who” system for consciousness of action and self-consciousness. *Consciousness and Cognition*, 7, 465–477.
- Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, 5, 382–385.
- Haggard, P., Martin, F., Taylor-Clarke, M., Jeannerod, M., & Franck, N. (2003). Awareness of action in schizophrenia. *NeuroReport*, 14, 1081–1085.
- Hagmayer, Y., Sloman, S. A., Lagnado, D. A., & Waldmann, M. R. (2007). Causal reasoning through intervention. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 86–100). Oxford: Oxford University Press.
- Jeannerod, M. (1999). The 25th Bartlett lecture. To act or not to act: Perspectives on the representation of action. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 52A, 1–29.
- Jeannerod, M. (2003). The mechanism of self-recognition in humans. *Behavioural Brain Research*, 142, 1–15.
- Jeannerod, M. (2009). The sense of agency and its disturbances in schizophrenia: A reappraisal. *Experimental Brain Research*, 192, 527–532.
- Knoblich, G., & Flach, R. (2003). Action identity: Evidence from self-recognition, prediction, and coordination. *Consciousness and Cognition*, 12, 620–632.
- Knoblich, G., & Kircher, T. (2004). Deceiving oneself about being in control: Conscious detection of changes in visuo-motor coupling. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 657–666.
- Knoblich, G., Stottmeister, F., & Kircher, T. (2004). Self-monitoring in patients with schizophrenia. *Psychological Medicine*, 34, 1561–1569.
- Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How we track time-varying events. *Psychological Review*, 106, 119–159.
- Leube, D., Knoblich, G., Erb, M., Grodd, W., Bartels, M., & Kircher, T. T. J. (2003). The neural correlates of perceiving one's own movements. *NeuroImage*, 20, 2084–2090.
- Lindner, A., Thier, P., Kircher, T. T. J., Haarmeier, T., & Leube, D. T. (2005). Disorders of agency in schizophrenia correlate with an inability to compensate for the sensory consequences of actions. *Current Biology*, 15, 1119–1124.
- Metcalfe, J., & Greene, M. J. (2007). Metacognition of agency. *Journal of Experimental Psychology: General*, 136, 184–199.
- Pacherie, E. (2008). The phenomenology of action: A conceptual framework. *Cognition*, 107, 179–217.
- Proust, J. (2000). Awareness of agency: Three levels of analysis. In T. Metzinger (Ed.), *The Neural Correlates of Consciousness* (pp. 307–324). Cambridge, MA: MIT Press.
- Repp, B. H. (2002). Perception of timing is more context sensitive than sensorimotor synchronization. *Perception & Psychophysics*, 64, 703–716.
- Repp, B. H., & Doggett, R. (2007). Tapping to a very slow beat: A comparison of musicians and non-musicians. *Music Perception*, 24, 367–376.
- Repp, B. H., & Knoblich, G. (2004). Perceiving action identity: How pianists recognize their own performances. *Psychological Science*, 15, 604–609.
- Repp, B. H., & Knoblich, G. (2007). Toward a psychophysics of agency: Detecting gain and loss of control over auditory action effects. *Journal of Experimental Psychology: Human Perception and Performance*, 33, 469–482.
- Ruby, P., & Decety, J. (2001). Effect of subjective perspective taking during simulation of action: A PET investigation of agency. *Nature Neuroscience*, 4, 546–550.
- Sato, A., & Yasuda, A. (2005). Illusion of sense of self-agency: Discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition*, 94, 241–255.
- Sebanz, N., & Lackner, U. (2007). Who's calling the shots? Intentional content and feelings of control. *Consciousness and Cognition*, 16, 859–876.
- Tsakiris, M., & Haggard, P. (2003). Awareness of somatic events associated with a voluntary action. *Experimental Brain Research*, 149, 439–446.
- Tsakiris, M., & Haggard, P. (2005). The rubber hand illusion revisited: Visuotactile integration and self-attribution. *Journal of Experimental Psychology: Human Perception and Performance*, 31, 80–91.
- Tsakiris, M., Haggard, P., Franck, N., Mainy, N., & Sirigu, A. (2005). A specific role for efferent information in self-recognition. *Cognition*, 96, 215–231.
- Tsakiris, M., Prabhu, G., & Haggard, P. (2006). Having a body versus moving your body: How agency structures body ownership. *Consciousness and Cognition*, 15, 423–432.
- van den Bos, E., & Jeannerod, M. (2002). Sense of body and sense of action both contribute to self-recognition. *Cognition*, 85, 177–187.
- Vignemont, F., & Fournier, P. (2004). The sense of agency: A philosophical and empirical review of the “Who” system. *Consciousness and Cognition*, 13, 1–19.
- Vignemont, F., Tsakiris, M., & Haggard, P. (2006). Body mereology. In G. Knoblich, I. Thornton, M. Grosjean, & M. Shiffrar (Eds.), *Human body perception from the inside out*. New York: Oxford University Press.
- Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.
- Wegner, D. M., Fuller, V. A., & Sparrow, B. (2003). Clever hands: Uncontrolled intelligence in facilitated communication. *Journal of Personality and Social Psychology*, 85, 5–19.
- Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation: Sources of the experience of will. *American Psychologist*, 54, 480–492.
- Wing, A. M. (1977). Perturbations of auditory feedback delay and the timing of movement. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 175–186.