Research

Judgements of agency in schizophrenia: an impairment in autonoetic metacognition

Janet Metcalfe[1,*], Jared X. Van Snellenberg[1,2], Pamela DeRosse[3,4], Peter Balsam[2,5] and Anil K. Malhotra[3,4]

[1]*Department of Psychology, Columbia University, 1190 Amsterdam Avenue, New York, NY 10027, USA*
[2]*Division of Cognitive Neuroscience, New York State Psychiatric Institute, 1051 Riverside Drive, New York, NY 10032-1098, USA*
[3]*Division of Psychiatry Research, The Zucker Hillside Hospital, 7559 263rd Street, Glen Oaks, NY 11004, USA*
[4]*Center for Psychiatric Neuroscience, Feinstein Institute for Medical Research, 350 Community Drive, Manhassat, NY 11030, USA*
[5]*Department of Psychology, Barnard College, 3009 Broadway, New York, NY 10027, USA*

We investigated judgements of agency in participants with schizophrenia and healthy controls. Participants engaged in a computer game in which they attempted to touch downward falling Xs and avoid touching Os. On some trials, participants were objectively in perfect control. On other trials, they were objectively not in complete control because the movement of the cursor on the screen was distorted with respect to the position of the mouse by random noise (turbulence), or it was lagged by 250 or 500 ms. Participants made metacognitive judgements of agency as well as judgements of performance. Control participants' judgements of agency were affected by the turbulence and lag variables—indicating that they knew they were objectively not in control in those conditions, and they were also influenced by their assessments of performance. The patients also used their assessments of performance but neither turbulence nor lag affected their judgements of agency. This indicated an impairment in agency monitoring. The patients, unlike the healthy controls, used only publically available external cues about performance in making judgements of 'agency' and did not rely on any additional access to internal self-relevant cues that were diagnostic in indicating whether or not they were, in fact, in control.

Keywords: metacognition; agency; schizophrenia

1. INTRODUCTION

The question of how an individual is able to determine whether it was the self or an alternative cause that was responsible for an action—metacognition of agency—is the concern of the present paper. This ability is crucial for learning and understanding one's own causal effect on the world, for all social interactions, and especially for coordination of individual and joint action (where the allocation of effort depends on knowing what one is doing and what the other is doing, and titrating one's own actions to accommodate those of others). This metacognitive capacity also underlies higher-order social judgements such as those that are necessary for the assignment of credit and blame. Understanding how people make judgements of agency and how other metacognitive judgements relate to these self-referential judgements is important in many domains. But, not all people make these judgements in the same way, and some have great difficulty in doing so accurately. In particular, the inability to keep the self straight—to know

what is self-produced and what is externally produced—characterizes a large part of the core deficit in patients with schizophrenia. Investigation of the cues that are used to make these self-relevant judgements as well as specification of the cues that patients with schizophrenia are unable to recruit may increase our understanding both of schizophrenia and of the processes underlying how people know about their own agency.

Following Tulving [1], Metcalfe & Son [2] have argued that there are three levels of metacognitive judgements: anoetic judgements (which are judgements about objects or events currently present in the world), noetic judgements (which are judgements concerned with internal representations, but without self-relevance) and autonoetic judgements (which are self-knowing judgements in which reference to the individual's self is implicated). While many researchers have argued that a central reason for studying metacognition is that it is the hallmark of human self-reflective consciousness, this characteristic only applies to autonoetic metacognition, and not to the other kinds. Reflection upon the self is not involved in either noetic or anoetic metacognition. Anoetic metacognition involves a judgement about a stimulus that is present at the time of judgement. It is sometimes

* Author for correspondence (jm348@columbia.edu).

thought that it is not even metacognition proper as no internal representation, or cognition, need be involved [3]. Most animals are capable of anoetic 'metacognition'. And although noetic metacognition, in which a judgement is made about an internal representation, is thought by virtually all researchers to really be metacognition (and some non-human animals have this capability; see the study of Smith *et al.* [4], for discussion of animal metacognition), it does not necessarily implicate *self-reflective* consciousness. No *self* need be involved. Judgements of agency, though, are truly autonoetic metacognitive judgements, being both self-reflective and self-knowing. They are a reflection on a cognition concerning the extent of one's own personal involvement and responsibility for an action. Isolating the cues that people use to accurately make these particular self-relevant judgements concerning how they know they are the agent, is, then, of specific interest for understanding the nature of human self-reflective consciousness. People with schizophrenia frequently have difficulty with attributions of just this sort.

Jeannerod [5] and others [6,7] have noted that patients with positive symptoms of schizophrenia, such as hearing voices and experiencing hallucinations and delusions, have difficulty in accurately reflecting upon their own agency. Such symptoms are also related to imaging findings showing hyperactivity in areas of the brain, in particular, the temporal parietal junction [8], that relate to the detection of a discrepancy between one's own intentions and the outcome that ensues [9]. These brain activity differences almost certainly relate to impairments in action monitoring. There are also known deficits in such patients in frontal brain areas that are associated with self-relevant processing [10–14], which likely relate to deficits in metacognitive judgements. With such impairments, an individual might easily make the mistake either of thinking that one's own internal thoughts came from outside and were produced by someone else rather than by oneself, or of believing that one was controlling events that were externally caused. Whether the representation one perceives came about because of one's own thought- or image-generation processes or was externally produced, is, at base, an attribution of agency, that is, a judgement about who or what was causal in producing the percept. These and other kinds of thought processes and inferences associated with schizophrenia [15–17] might well result from impairments in a circuit that normally, accurately and efficiently, evaluates agency.

Although healthy adults are usually able to make accurate judgements of agency [18], even they can sometimes be fooled about whether or not they were the agent [19,20]. Furthermore, people at different stages of development make judgements of agency that are systematically sensitive to different parameters [21]. The findings of illusions of agency, and of systematic differences in these judgements even in healthy adult populations, suggest that there are a number of distinct cues that contribute to agency judgements. Both the cues and the judgement processes appear to be malleable.

The idea that metacognitive judgements of agency are based on cues, rather than direct knowledge [22,23], is consistent with the widely held view that all metacognitive judgements are cue-based. There are many cases, detailed in the voluminous metacognitive literature, in which it has been shown that certain judgements rely on different cues from one another (see [24–30] for discussion and evidence concerning the cue-based nature of different metacognitive judgements). Understanding which cues are used for making judgements of agency, as well as what the neural circuitry is that underlies each of them, is important in ameliorating distortions seen in these judgements in people with schizophrenia. Studying the locus of the deficit in patients who have impairments in this particular metacognitive domain may also allow more intensive scrutiny of the cues and mechanisms contributing to these central metacognitive judgements in healthy people. The investigation of metacognition of agency, though relatively new, points to four cues, or sources of information, that appear to contribute to these judgements. Interestingly, while the judgements themselves are concerned with whether the self was or was not responsible for an action, not all of the cues used to make these judgements are internal self-relevant cues.

2. CUES CONTRIBUTING TO JUDGEMENTS OF AGENCY

(a) *Judgements of performance*

Perhaps surprisingly, the single most important factor that has emerged as a predictor of people's agency judgements is another metacognitive judgement, namely, judgements of performance. While judgements of agency are autonoetic—being explicitly about the role of the self in an action—judgements of performance need not reference or even reflect the self. They are merely noetic (i.e. judgements about a representation, but without the necessary involvement of the self that would make them autonoetic). In the task that we will investigate [18] people play a computer-based game of having a cursor touch Xs and avoid descending Os by moving the computer mouse. At the end of each trial, they are asked for a judgement of performance. The judgement of performance does not, itself, require that the individual participant be the agent. Such a judgement about the proportion of Xs touched and Os avoided on the last trial could be made even if someone other than the subject had been controlling the mouse. In short, this judgement is noetic, not self-referential, and need not imply access to the participant's own role as the person controlling the mouse to touch the Xs. Even though this assessment says nothing about who was responsible for the action, people's perception of performance is, nevertheless, an important cue used to make judgements of agency: when performance is perceived to be good, agency is claimed; when performance is perceived to be poor, agency is denied.

Regression analyses directed at determining the sources of information that contribute to normal adults' metacognition of agency have revealed that people's perception of their level of success on the task on each trial is a strong contributor [31]. The self-relevant autonoetic agency judgement, then, is based in large part on a non-self-referential noetic

judgement concerning the goodness of performance. It is not necessary to be the agent at all, or to evaluate any internal or visceral cue to which one has privileged access, to make a judgement of performance. And yet people, apparently, use these judgements to a large but, importantly, not exclusive extent, to evaluate whether or not they were in control of an action.

While acknowledging that people do use their judgements of performance to make agency judgements, Metcalfe and co-workers [18,21] have argued that since these performance judgements do not indicate the *source* of the action, they should be factored out of the analyses, to allow investigation of whether people were sensitive, in a *veridical way,* to cues implying that the self was or was not the agent. They also noted that people's judgements of agency should be assessed relative to their *perception* of their performance, rather than their actual performance, as it is not how the person is doing objectively that counts, but rather how they think they are doing. The use of judgement of performance, to anchor people's judgements of agency, also provides some leverage on how individuals use the rating scales. The question of interest, in evaluating the accuracy of people's judgements of agency, then, is whether—in the conditions in which they are not fully in control—they pick up on their lack of control, over and above their perception of their overt performance. Thus, to evaluate people's metacognition of agency in past experiments [21,31], people's judgements of agency were compared with their judgements of performance. Because there may be scaling effects in how people ground the judgement of agency scale, the 'control' condition—in which there were no distortions of their actual control—was used to anchor their use of the performance and the agency scales. In the analyses that follow in this article, we, too, will use the difference between judgements of performance and judgements of agency, and will use the control condition as the baseline against which to evaluate differences in these two judgements that occur in the experimental conditions in which objective control was distorted.

Before leaving the topic of the role of judgements of performance on judgements of agency, it is notable that other researchers have also shown that noetic metacognitive and autonoetic judgements sometimes appear to be intertwined. Cosentino *et al.* [32] have shown that noetic judgements of learning are strongly related to autonoetic judgements of agency. Similarly, David *et al.* [33] have discussed the relation of metacognition to anosagnosia, and Cosentino *et al.* [34] have shown that inaccuracy in metamemory judgements (again, noetic judgements) are associated with a lack of awareness of memory deficit (anosagnosia)—an awareness of one's own capabilities that would seem to involve self-knowing consciousness. And, finally, Fleming *et al.* ([35], and see [36]) showed that the same brain area (BA10) that Miele *et al.* [31] isolated as being more strongly activated in making self-relevant judgements of agency as contrasted with judgements of performance, is also, itself, implicated in noetic metacognitive judgements. Indeed, as Fleming *et al.* [35] showed, individual differences in the accuracy of (noetic) metacognitive

judgemknts were shown to be related to structural brain differences in this area. Thus, the two kinds of judgements—while conceptually distinct—may be functionally related at a deep level.

(b) Discrepancy monitoring of the correspondence between plan and outcome

Frith *et al.* ([15], and see [37−40]) have proposed a brain-based framework for motor control that relates in a natural way to people's metacognitions of agency. The 'comparator' model was originally devised to explain how people make fine-grained corrections of motor movements, and has been shown to be valuable in illuminating one source of information that could provide a focal cue in agency judgements. According to this scheme, when a person has a goal, it gives rise to an internal model of their intentions (inverse model) and expectations (forward model) about achieving the goal. This initiates a motor plan that provides the specifications about what needs to be done to achieve the goal. The plan or expectation runs off in real time simultaneously with the person's motor actions. A comparator mechanism evaluates the correspondence of the actions and the plan. A match between the expectation and the outcome indicates that the person's intentions corresponded to what happened, and no motor adjustment need occur. A discrepancy provides a signal to the motor system indicating that the movement needs to be adjusted to achieve the goal. The discrepancy can also be used by the metacognitive system as a cue indicating that something or someone else was interfering with the intended action: the person was not in complete control. For example, if turbulence or noise were introduced into the instrument the person is controlling, then the plan for the motor actions would fail to match what happened, because of the noise. In such a situation, the discrepancy may be a cue used in a judgement process that provides a reliable indicator that the person was not completely in control.

In schizophrenia, either the plan or the internal feedback from the person's own actions may be distorted, and this may give rise to misattributions of control [41] attributable to such a discrepancy monitoring mechanism. The model even points to components of a brain network (the temporal parietal junction, with cerebellum involvement) where one might seek to find evidence of this discrepancy. Given that discrepancy detection is closely linked to motor control, one might expect to see altered motor control in patients whose impaired metacognition of agency is due to an impairment in the forward model. For example, Synofzik *et al.* [6] showed that patients with positive symptoms showed a higher threshold for detecting discrepancies in feedback rotations, indicating an impairment in the precision of sensory predictions. Additionally, many patients with schizophrenia exhibit motor impairments, as well as abnormalities indicating irregularities in this action-monitoring system.

However, not all patients exhibit such motor impairments. Knoblich *et al.* [42] conducted an experiment in which the participants—both healthy controls and people with schizophrenia—attempted to keep their stylus on a circling moving dot on the screen. When

the dot accelerated off course, such that the participant had to change their action pattern to allow it to continue following the correct path, participants were supposed to keep the dot on the circle and to indicate that they detected the distortion that was occurring. People with schizophrenia were able to do the motor task—altering their motor behaviour—as well as the healthy control participants. However, they were much slower and less likely to consciously *notice* the distortion (i.e. to have metacognition about the change) than were healthy controls. Thus, it appears that the metacognitive assessment can sometimes be dissociated from the motor aspects of the task in schizophrenia, suggesting that the metacognitive judgement processes themselves may be independent of the action-monitoring guiding motor performance.

(c) *Reward*

While it is logically possible that the feeling of being in control is just a lack of feeling out of control [43], it is also possible that positive feelings of agency are, themselves, neurally coded, and distinct from such a proposed default state of not being out of control. Feelings of being *in control* have been claimed by a number of 'positive' psychologists [44] who stress the role of self-determination, to be both intrinsically rewarding, and to be associated with learning. Consistent with this notion, in Miele *et al.*'s [31] fMRI study, it was found that trials in which participants reported a high level of feeling 'in control' were associated with increased activity in the presupplementary motor areas, the rostral cingulate zone and the dorsal striatum, regions that are linked to self-initiated action and reward. The activation of this intention and reward-related system, in conjunction with feelings of being in control, rather than, say, deactivation of the temporal parietal junction area (which would indicate a default state of *not* being *out of* control) lends some credibility to the idea that feeling 'in control' may, itself, be a separable state with consequences.

Kirkpatrick *et al.*'s [45] work also converges on the idea that the reward system is related to positive agency judgements. In their study, methamphetamine users, after receiving either the drug or a placebo, engaged in a motor agency task similar to the one used in this article. Insofar as methamphetamine has its effects on the dopamine/reward system, effects of the drug itself on the judgements of agency might be thought to be mediated by the reward system. Interestingly, then, although there was no difference in performance on the task depending upon whether the participants were on methamphetamine or not, their judgements of agency were increased, under conditions of objectively perfect control, when they were on the drug rather than on placebo. The drug, evidently, made them feel more agentic, or more in control.

Finally, Tricomi *et al.* [46] have found that reward-related areas were activated during conditioned learning, but only when the participants were aware of the contingency between their button presses and the outcomes. Being aware of the contingency between one's actions and the outcome, of course, could be rephrased as knowing that one was in control. If we

interpret the results in this way, they would suggest that learning, associated with activation in the striatum, is related to feelings of agency. These data, then, suggest that knowledge of agency may be necessary for reward-related learning.

The potential involvement of the reward system in metacognition of agency may be of importance in schizophrenia because of the involvement of dopamine in schizophrenia. It is possible, perhaps even likely, that people with schizophrenia have abnormal responses to reward that relate in a complex way to misperceptions of agency. The role of reward and its impact upon people's metacognitions of being in control may, therefore, have special interest in this context.

(d) *Temporal delay*

A judgement of agency is a special case of a judgement of causality, in which the question is whether the *self* is the causal agent. It would, therefore, be expected that factors affecting people's perception of causality would also affect judgements of agency. Perhaps the most studied of these factors that affect judgements of causality is temporal contiguity. Michotte ([47], and see [48]) has shown that when one moving object makes contact with another, and then the second, without any delay, begins to move, this interaction is perceived as causal with the first object causing the movement of the second. Michotte called this phenomenon the 'launching' effect. The perception of causality is systematically diminished as a lag is interposed between the movement of A and the movement of B. It follows that feelings of agency should also be decreased if a delay is interposed between one's act and the result.

In keeping with this idea, Blakemore *et al.* [49] have shown that when there is no temporal delay between the act of attempted self-tickling and the resultant self-stimulation, healthy individuals are unable to tickle themselves. They argue that the reason an individual cannot tickle him/herself is that the concordance between plan and outcome results in a diminution in the perceived stimulation. No such diminution occurs with a mismatch, and the tickle sensation is, hence, perceived when another person is responsible for the tickling. However, when a delay is interposed between the act of tickling oneself and the resultant self-stimulation (by means of a mechanical device), healthy individuals can self-tickle, underlining the role of temporal delay. It is notable, in this context, that Blakemore *et al.* [50] found that, unlike healthy participants, patients with schizophrenia were able to tickle themselves even without a temporal delay being interposed.

The data of Schlottman & Shanks [51] show systematic decreases in causality ratings as delay is increased. Nevertheless, even at fairly large delays people still judged A and B to have a causal relation, consistent with the Kantian idea that causality is inferred as long as there is *any* rule that is seen to mediate between A and B. In the experiment below, we equated the amount of discrepancy between the position of the cursor and the position of the mouse in a pure 'turbulence' condition, in which no rule mediates, and in a time-delayed condition where there was a mediating rule. Past research has indicated

that healthy adult participants feel less out of control in the time-delayed condition, which has a mediating rule, than in the turbulence condition which does not [21]. People with schizophrenia, though, may have difficulty picking up on such a subtle mediating rule, and hence may not use this cue.

(e) *The judgement process*

Finally, while the cues used and the sensitivity towards them may vary from person to person, and some or all of them may be impaired in people with schizophrenia, it is possible that these cues to agency could all be veridical, and yet an impairment in metacognition of agency could still result. It is possible that the judgement process itself could be distorted. An fMRI study has shown that there is a difference in neural processing in anterior prefrontal cortex between making a judgement of agency as contrasted to making a judgement of performance [31]. In other research, this area has been shown to be associated with other kinds of self-referential processing [12,52,53] and metacognitive judgements [34,36]. The self-referential metacognitive judgement appears to be distinctive. It is possible that patients could have either intact or impaired ability to make such self-referential judgements. However, if this judgement process were impaired, agency judgements would be expected to be impacted even in the presence of veridical cues.

3. EXPERIMENT

The task employed was the same as has been used in past experiments [21] in which metacognition of agency was compared between young adults, children and elders. As mentioned above, participants played a computer game in which they moved the mouse to touch downward falling Xs on the screen and, at the same time, to avoid touching Os. Objective control of the cursor by the mouse could be undistorted, in the control condition (i.e. the person was objectively in full control), or could be altered by means of a lag in the relation between the mouse position and the cursor position or by turbulence (random noise) intervening between the mouse position and the cursor position. At the end of each trial, the participant made a judgement of his/her own control, that is a judgement of agency, as well as a judgement of performance. This task allowed us to investigate whether manipulations that objectively altered the person's control were open to accurate metacognitive assessment.

4. METHOD

(a) *Participants*

The patient group included 22 patients recruited from the Zucker Hillside Hospital (ZHH), a division of the North Shore-Long Island Jewish Health System (NSLIJHS), in Glen Oaks, NY to a protocol designed to assess functional disability in stable outpatients. Inclusion criteria for patients included clinical stability as defined by no hospitalization in the last six months, between 18 and 59 years of age with a DSM-IV diagnosis of schizophrenia or schizoaffective disorder, and no substance abuse in the preceding one month. All patients were on antipsychotic medication at the time of testing. The mean age of the patient sample was 42.3 years (s.d. = 11.1) and 40.9% were female. Healthy comparison subjects were recruited from the general population through the ZHH Healthy Control Initiative. Potential controls were excluded if they had a DSM-IV axis I diagnosis or a first-degree relative with a known or suspected axis I disorder. The mean age of the control sample was 38.1 years (s.d. = 11.3) and 45 per cent were female. Patients and controls with a history of central nervous system trauma, neurological disorder (including seizures), mental retardation or known genetic disorder were excluded.

(b) *Diagnostic measures*

Patients' diagnoses were established with the structured clinical interview for DSM-IV (SCID) [54] and confirmed by diagnostic consensus conference, which uses expert clinical opinion alongside SCID and corroborating medical record information. Brief psychiatric rating scale (BPRS) mean was 27.2 (5.8) and the scale for the assessment of negative symptoms (SANS) was 29.3 (12.2). Comparison subjects were assessed with the SCID–non-patient edition to rule out axis I diagnoses.

(c) *Apparatus*

All experiments were conducted on individual iMac computers, used with a mouse, and mouse pad. Participants were tested individually.

(d) *Procedure*

The instructions were: 'Throughout this experiment you are going to play a game in which you will use the computer mouse to move a box on a grey track. Your job is to touch all of the Xs as they come into range and to avoid touching any of the Os. After each trial, you will be asked to assess your performance. If you felt you got all of the Xs, and avoided all of the Os, you should click to the far right of the blue bar, indicating everything correct. If you felt you got none of the Xs, and touched all of the Os, then you should click to the far left, indicating nothing correct. You may also click anywhere in between. You will also be asked to assess how in control you felt. If you felt you were in complete control, click to the far right of the red bar. If you felt that you had no control, click to the far left. You may also click anywhere in between'.

In this experiment, the performance judgement was always made before the judgement of agency. The constant order was used to minimize possible confusion. Previous experiments that have used either only an agency judgement or only a performance judgement on each trial [18,31] have produced comparable results with those that have used both judgements on every trial [21].

Participants practised both playing the game and making judgements, under the supervision of the experimenter, who made sure that the participant understood how the task and how the rating scales worked by having the participant report what each judgement meant, following each practice trial. The practice trials were repeated as many times as was

necessary. After the practice trial(s), the experimenter asked if there were any questions, and if there were, he or she answered them. At the end of the experiment, the participant was questioned about what he or she had done, and was paid and thanked for participating.

(e) Design

The experiment included six within-participant conditions: a control condition in which the participant had perfect control of the mouse, a short lag condition (Lag1) in which the cursor responsiveness lagged the mouse position by 250 ms, a long lag (Lag2) condition in which the cursor position lagged the mouse position by 500 ms, a small amount of turbulence (Turb1) condition, which was discrepancy matched (as will be described shortly) to the Lag1 condition, a large amount of turbulence (Turb2) condition, which was discrepancy matched to the Lag2 condition and a 'magic' condition, which artificially inflated performance, and was important so that participants did not become discouraged. The magic condition showed no differences between the control and patient participants on any measure, and is, therefore, not included in the analyses or discussed further. There were four replications of all conditions.

The amount of noise in the turbulence conditions was matched with the amount of discrepancy between the mouse position and the cursor position in the lag conditions. This was done by measuring every 8 ms, on the first lag trial, the discrepancy between the mouse position and the cursor position, and then re-randomizing these signed difference scores and adding them to the cursor position at each 8 ms interval in the appropriate turbulence condition. This added noise was smoothed to prevent sudden jerks. Because of this matching algorithm, the amount of discrepancy—where discrepancy is considered to be the difference between the mouse position and the cursor position at each sampled position over the entire 15 s trial—was the same in the lag condition and in the matched turbulence condition, and so type of discrepancy and amount of discrepancy could be treated as factors. The difference between the two types of discrepancy conditions was that the discrepancy between the mouse position and the cursor position in the turbulence condition was random, while in the lag condition, it was lawfully mediated by a time lag rule: if one were to shift the cursor position function back by 250 or 500 ms, in the lag conditions, it would match the mouse position function perfectly. The lag and turbulence conditions, with high and low levels of discrepancy, therefore, comprised a 2×2 design. In the control condition, which is used as a baseline, there was no discrepancy between the cursor position and the mouse position.

To equate the discrepancy as outlined above, the lag condition had to come first, which constrained the randomization of the order of conditions within block, though all conditions were well distributed over the entire session. The data were, therefore, analysed both with and without the first lag trial. Because there was no difference depending on its inclusion, it was included in the analyses that are reported below.

The data from the four trials in each condition for each participant were collapsed.

The two metacognitive-dependent variables of central interest were people's judgement of performance (i.e. how well did they think they had done on touching the Xs and avoiding the Os) and their judgement of agency (i.e. how in control did they think they were). Both were measured on an analogue scale coded from 0 to 1.0. We also computed performance using hit rate (i.e. the proportion of times the person touched in-range Xs) and false alarm rate (i.e. the proportion of times the person, incorrectly, touched Os). In past experiments on this paradigm, as in the present experiment, hit rate and d' (a measure of goodness of performance) were highly correlated, and only the former has shown a strong relation to people's judgements of performance, with false alarm rate having only a very small impact on their judgements [18]. We, therefore, report only hit rate here.

5. RESULTS

In the results that follow, in cases where a participant did not finish all trials, their data are included as long as they completed at least two trials in each condition. A value of $p < 0.05$ was used to determine significance.

(a) Performance

As shown in figure 1, there was a main effect of condition on hit rate, $F_{4,160} = 136.04$, $p < 0.01$. There was also a main effect of group on hit rate, $F_{1,40} = 6.15$, $p < 0.02$ but this effect was qualified by significant interaction between condition and group, $F_{4,160} = 6.27$, $p < 0.01$. Post hoc tests showed that the healthy controls performed significantly better than did the patients only in the control condition ($t_{40} = 3.43$, $p < 0.01$).

(b) Metacognition of performance

Figure 1 also shows that judgements of performance closely tracked hit rate in both groups. There were strong correlations between hit rate and judgements of performance (collapsing across conditions, within participants and using Fisher's r-to-Z transformation throughout to normalize the distributions). The mean correlation ($\pm$ s.d.) for control participants was 0.87 ± 0.48, which was significantly greater than zero ($t_{19} = 11.52$, $p < 0.01$). For patients, the mean correlation was 0.64 ± 0.43, which was also significantly greater than zero ($t_{21} = 7.81$, $p < 0.01$). Although the correlation for controls was significantly greater than that for patients ($t_{40} = 3.81$, $p < 0.01$), the correlations shown between performance and judgements of performance by the patients were still very high and comparable with the correlations found, in this same paradigm, with children ($r = 0.67$), and elders ($r = 0.81$ [21])—groups that showed very good metacognition of agency.

A measure of calibration for each participant in each condition was computed based on the difference between their hit rate and their judgement of performance. As can be seen from figure 1, judgements of performance were slightly lower than hit rate for the
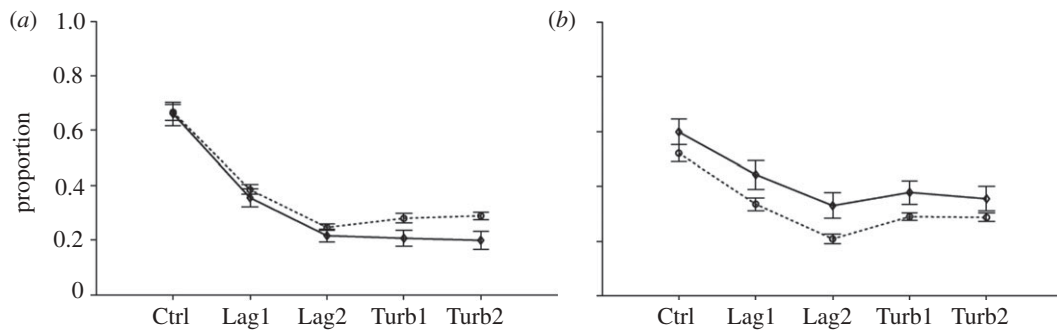
Figure 1. (*a*) Hit (circles and dotted lines) rates and judgements of performance (diamonds and solid lines) in control participants, and (*b*) patients with schizophrenia.
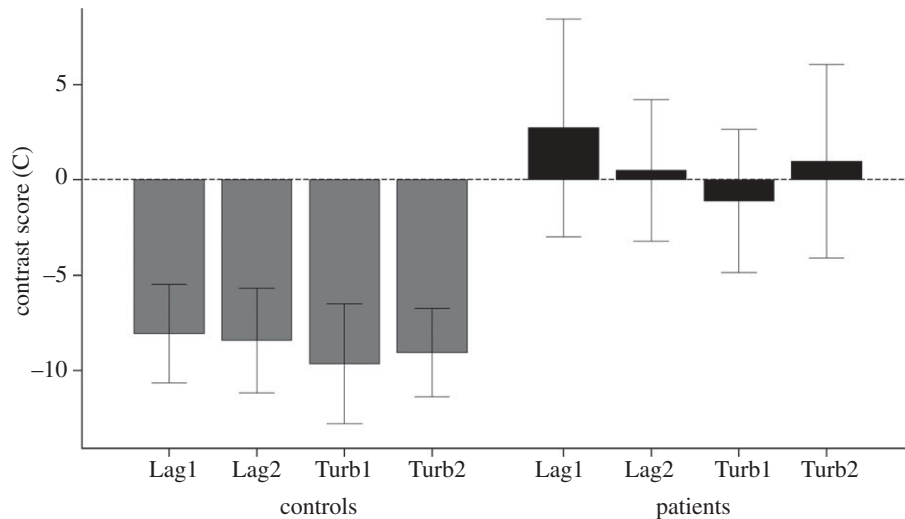

Figure 2. Contrast scores in the turbulence and lag conditions for control participants, and for patients with schizophrenia. A negative score indicates that the person appropriately realized that the manipulation had distorted their control.

control participants, whereas judgements of performance were higher than performance for the patients. Statistically, while there was neither a main effect of condition nor an interaction between condition and group, there was a significant calibration main effect of group ($F_{1,40} = 9.72$, $p < 0.01$). This difference in calibration between groups—showing that the healthy controls were slightly underconfident while the patients were overconfident—may relate to a 'reward'-related difference in perception between the two groups: the patients, but not the controls, thought they had done better than they had.

In summary, then, both the patients' actual performance, and their noetic metacognition, as measured by the correlation between their judgements of performance and their performance, were well above chance, though not as good as those of the healthy control participants.

(c) *Metacognition of agency*
We next asked whether participants picked up on their lack of control, appropriately, over and above their perception of their overt performance. To evaluate whether people experienced a greater decrement in their feelings of agency in the turbulence and lag conditions, we computed summary 'agency' scores, namely, the contrast: (judgement of performance$_C$ – judgement of agency$_C$) – (judgement of performance$_E$ – judgement of agency$_E$) where the subscript C refers to the control

condition and E refers to either Turb1, Turb2, Lag1 or Lag2.[1] This summary score should be negative, so long as people realized that their performance, in the experimental condition being considered, was not entirely due to their own control. A zero means that they thought they were in control.

As can be seen from figure 2, the control participants were sensitive to both the turbulence and the lag conditions' effect on decreasing their control—they showed strongly negative contrast scores. In contrast, the patients were insensitive to the manipulations—showing contrast scores of near zero. There was a significant main effect of group, $F_{1,40} = 4.55$, $p < 0.05$, but no other main effects or interactions. Furthermore, one-sample t-tests revealed that control participants had significantly negative contrast scores in all four conditions (all $ps < 0.01$), while patients' contrast scores were not different from zero in any of the conditions (all $ps > 0.66$). These patient data reveal a lack of metacognition of agency unlike that seen in any group that we have studied to date. In contrast to the data presented here, all previous groups tested on this paradigm have shown negative values on all four contrast scores.

Although numerically the controls showed slightly more negative scores in the turbulence conditions than in the lag conditions (as has been shown with young adults in past experiments, [21]), the interaction was not significant.
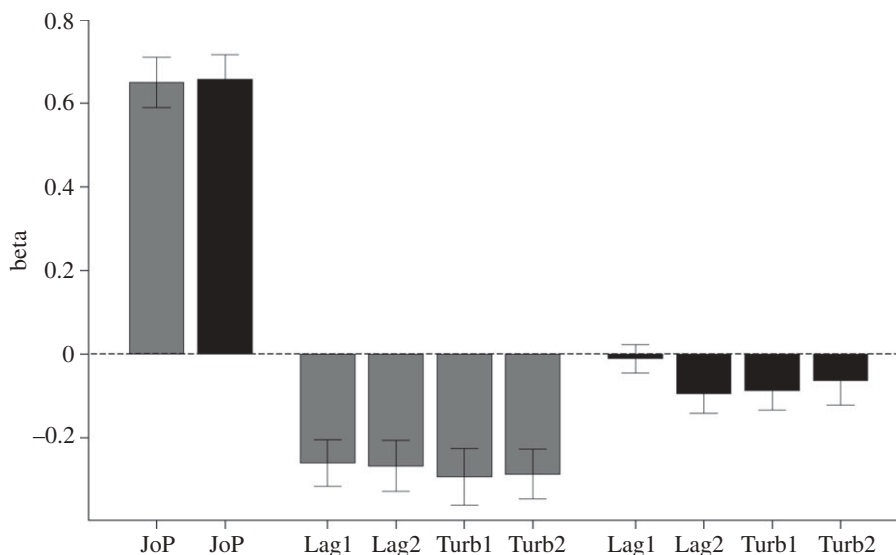
Figure 3. Cues used by control (grey bars) participants and patients with schizophrenia (black bars) to make judgements of agency. The figure shows mean betas from regression models calculated for each participant, with individual trials as the unit of analysis, and in which judgement of agency was the criterion variable. The potential predictors were judgement of performance (JoP) and the experimental conditions—Lag 1, Lag 2, Turb1 and Turb2.

We conducted correlation analyses to determine which, if any, symptoms in the patients' diagnostic profiles predicted the above contrast scores. Insofar as negative contrast scores indicated sparing of metacognition of agency, we hypothesized that some symptoms, or some lack of symptoms, might predict such sparing. However, the results of these analyses failed to show significant results selective to any symptoms. We conducted a similar analysis with the raw judgement of agency scores, and, again found no correlation between particular symptoms and scores.

Finally, we conducted a regression analysis to investigate what information contributed to participants' judgements of agency. As can be seen from the (normalized) beta values given in figure 3, the control participants' judgements of agency were predicted by their judgements of performance as well as by the turbulence and lag conditions in which control was objectively impaired. In contrast, the patients' judgements of agency were predicted *only* by their judgements of performance. Control participants' judgements of agency were influenced by their judgements of performance ($t_{19} = 11.37$, $p < 0.01$), as were those of patients ($t_{21} = 11.03$, $p < 0.01$), and the two groups were not different ($t_{40} = 1.02$, $p = 0.32$). However, with all other predictors, there was a difference between the controls and the patients. Control participants' judgements of agency were significantly influenced by the Lag1 condition ($t_{19} = 3.87$, $p < 0.01$), by the Lag2 condition ($t_{19} = 4.07$, $p < 0.01$), by the Turb1 condition ($t_{19} = 4.02$, $p < 0.01$) and by the Turb2 condition ($t_{19} = 4.31$, $p < 0.01$). In contrast, as can be seen from the figure, the patients' judgements of agency were not significantly influenced by any of these conditions. As might be expected, in each of the four cases, the influence of the condition on judgement of agency was significantly greater for the control participants than for the patients (respectively, for Lag1, Lag2, Turb1 and Turb2: $t_{40} = 3.69$, $p < 0.01$; $t_{40} = 3.04$, $p < 0.01$; $t_{40} = 2.94$, $p < 0.01$; $t_{40} = 2.98$, $p < 0.01$.)

6. DISCUSSION
The results presented here provide further indication that people's metacognition of agency is based on specific cues that are evaluated by a judgement process. The results also provide support for the separation of noetic and autonoetic metacognition. The healthy controls used both noetic (performance-related, that could be purely external) and autonoetic (internal) cues in making their agency judgements. These data indicate, however, that the patients used only the noetic cues, and did not recruit the autonoetic cues in making their judgements of agency.

The patients with schizophrenia performed very well on many aspects of the task. Moreover, their noetic metacognition, as given by the high correspondence between their judgements of performance and their actual performance, was good, though not quite as good as that of healthy controls. Thus, they did not show a profound deficit in all kinds of processing, or even in all kinds of metacognitive processing. However, they showed no sensitivity whatsoever to internal factors that objectively provide the kind of cues that healthy controls use to determine, accurately, when they are in control and when they are not. Unlike healthy control participants, the patients with schizophrenia appeared to be unaware of the presence of turbulence in the mouse controls, or the fact that the response of the cursor was altered by a time lag of up to half a second. Healthy control participants know, very reliably, that they are 'out of control' under those circumstances.

The patients were not random in making their judgements of agency. The regression analyses showed that they did use one cue that is also used by healthy control subjects, namely the perceived goodness of performance. Furthermore, they appear to use this cue to about the same extent as do the healthy controls. But this was the *only* cue that the patients with schizophrenia used. Judgements of agency were, apparently, made without evaluation of any internal or visceral cues, or, indeed without any reference to the self,

insofar as judgements of performance could be made purely externally and visually, by simply observing and remembering what happened in the trial. It was not necessary to know who was the agent (or, indeed, that there even was a human agent) to make such a judgement. Thus, the cues that provided the input for the purportedly self-referential judgements were, for the patients, lacking in any privileged or internal information. It is of some interest that Synofzik *et al.* [6] have observed a similar reliance on external visual cues in patients with schizophrenia. They propose that this reliance on external cues occurs because patients' internal cues are unreliable. The fact that basic noetic metacognitive processes—though not the self-referential ones—seem to be spared in the patient group, indicates, along with other evidence surveyed herein, that self-referential or autonoetic metacognition, while building on more basic metacognitive processes that are noetic in nature, may, nevertheless be both different and separable.

It is clear from the present study that some aspects of the positive symptoms of schizophrenia could arise from a deficit in the perception of agency based on either a difficulty in perceiving the autonoetic cues or on using those cues to make judgements. It would be interesting to know if this difficulty encompasses all autonoetic cues or only those related to discriminating the relation between actions and outcomes. Further, it might be possible to explicitly train patients to discriminate autonoetic cues, perhaps producing a significant reduction in some positive symptoms such as delusions. Current pharmacological treatments only temporarily meliorate positive symptoms. In contrast, interventions such as the one suggested above hold the potential for more permanent alterations by directly treating the underlying deficit.

Finally, the present results point to the interweaving of different kinds of metacognitive cues in the service of an externally posed task. This study asked for a judgement that directly focused on participants' own personal involvement as a causal agent. And yet, despite the task requirements, those judgements were made by using cues related to external outcomes that have no necessary connection to the role of the self in the action. Judgements about agency were, in healthy control participants, also based on internal cues indicating distorted objective control. These internal cues provided reliable information about the individual's role as an agent. In contrast, the patients used only the performance cues, and did not access the internal cues that could allow accurate evaluation of the causal role of the self in action. The dissociation between the judgements of patients and healthy controls provides support for the importance of a distinction between noetic and autonoetic metacognition.

All subjects provided written informed consent to a protocol approved by the NSLIJHS Institutional Review Board.

ENDNOTE

[1]In the control condition, the patients had agency judgements lower (58.60, s.d. = 21.63) than performance judgements (60.28, s.d. = 21.76), while the healthy controls' agency judgements were higher (69.79, s.d. = 22.22) than performance judgements (65.85, s.d. =

20.09). This latter pattern has been found in other studies with healthy participants.

REFERENCES

1 Tulving, E. 1985 Memory and consciousness. *Can. Psychol.* **26**, 1–12. (doi:10.1037/h0080017)
2 Metcalfe, J. & Son, L. K. In press. Anoetic, noetic, and autonoetic metacognition. In *The foundations of metacognition* (eds M. Beran, J. L. Brandl, J. Perner & J. Proust). Oxford, UK: Oxford University Press.
3 Metcalfe, J. & Kober, H. 2005 Self-reflective consciousness and the Projectable Self. In *The missing link in cognition: origins of self-reflective consciousness* (eds H. S. Terrace & J. Metcalfe), pp. 57–83. Oxford, UK: Oxford University Press.
4 Smith, J. D., Couchman, J. J. & Beran, M. J. 2012 The highs and lows of theoretical interpretation in animal-metacognition. *Phil. Trans. R. Soc. B* **367**, 1297–1309. (doi:10.1098/rstb.2011.0366)
5 Jeannerod, M. 2009 The sense of agency and its disturbances in schizophrenia: a reappraisal. *Exp. Brain Res.* **192**, 527–532. (doi:10.1007/s00221-008-1533-3)
6 Synofzik, M., Thier, P., Leube, D. T., Peter Schlotterbeck, P. & Lindner, A. 2010 Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions. *Brain* **133**, 262–271. (doi:10.1093/brain/awp291)
7 Voss, M., Moore, J., Hauser, M., Gallinat, J., Heinz, A. & Haggard, P. 2010 Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. *Brain* **133**, 3104–3112. (doi:10.1093/brain/awq152)
8 Decety, J. & Lamm, C. 2007 The role of the right temporo-parietal junction in social interaction: how low-level computational processes contribute to meta-cognition. *Neuroscientist* **13**, 580–593.
9 Farrer, C., Franck, N., Frith, C. D., Decety, J., Georgieff, N., d'Amato, T. & Jeannerod, M. 2004 Neural correlates of action attribution in schizophrenia. *Psychiat. Res.* **131**, 31–44. (doi:10.1016/j.pscychresns.2004.02.004)
10 Carter, C. S., MacDonald, A. W., Ross, L. L. & Stenger, V. A. 2001 Anterior cingulate cortex activity and impaired self-monitoring of performance in patients with schizophrenia: an event-related fMRI study. *Am. J. Psychiat.* **158**, 1423–1428. (doi:10.1176/appi.ajp.158.9.1423)
11 Flashman, L. A., McAllister, T. W., Johnson, S. C., Rick, J. H., Green, R. L. & Saykin, A. J. 2001 Specific frontal lobe subregions correlated with unawareness of illness in schizophrenia: a preliminary study. *J. Neuropsychiat. Clin. Neurosci.* **13**, 255–257. (doi:10.1176/appi.neuropsych.13.2.255)
12 Gilbert, S. J., Spengler, S., Simons, J. S., Steele, J. D., Lawrie, S. M., Frith, C. D. & Burgess, P. W. 2006 Functional specialization within rostral prefrontal cortex (area 10): a meta-analysis. *J. Cogn. Neurosci.* **18**, 932–948. (doi:10.1162/jocn.2006.18.6.932)
13 Juckel, G., Schlagenhauf, F., Koslowski, M., Wüstenberg, T., Villringer, A., Knutson, B., Wrase, J. & Heinz, A. 2006 Dysfunction of ventral striatal reward prediction in schizophrenia. *Neuroimage* **29**, 409–441. (doi:10.1016/j.neuroimage.2005.07.051)
14 Pomarol-Clotet, E. *et al.* 2010 Medial prefrontal cortex pathology in schizophrenia as revealed by convergent findings from multimodal imaging. *Mol. Psychiat.* **15**, 823–830. (doi:10.1038/mp.2009.146)
15 Frith, C. D., Blakemore, S. & Wolpert, D. M. 2000 Interactive report explaining the symptoms of

schizophrenia: abnormalities in the awareness of action. *Brain Res. Rev.* **31**, 357–363.

16 Jeannerod, M. 2009 The sense of agency and its disturbances in schizophrenia. *Exp. Brain Res.* **192**, 527–531. (doi:10.1007/s00221-008-1533-3)

17 Pacherie, E., Green, M. & Bayne, T. 2006 Phenomenology and delusions: who put the 'alien' in alien control? *Conscious. Cogn.* **15**, 566–577.

18 Metcalfe, J. & Greene, M. J. 2007 Metacognition of agency. *J. Exp. Psychol. Gen.* **136**, 184–199. (doi:10.1037/0096-3445.136.2.184)

19 Wegner, D. M. & Wheatley, T. 1999 Apparent mental causation: sources of the experience of will. *Am. Psychol.* **54**, 480–492. (doi:10.1037//0003-066X.54.7.480)

20 Wegner, D. M., Sparrow, B. & Winerman, L. 2004 Vicarious agency: experiencing control over the movements of others. *J. Pers. Soc. Psychol.* **86**, 838–848. (doi:10.1037/0022-3514.86.6.838)

21 Metcalfe, J., Eich, T. S. & Castel, A. 2010 Metacognition of agency across the lifespan. *Cognition* **116**, 267–282. (doi:10.1016/j.cognition.2010.05.009)

22 Carruthers, P. 2011 *The opacity of mind: an integrative theory of self-knowledge.* Oxford, UK: Oxford University Press.

23 Ryle, G. 1949 *The concept of mind.* London, UK: Hutchinson (page references are to the 2000 republication, London, UK: Penguin Books).

24 Benjamin, A. S., Bjork, R. B. & Schwartz, B. L. 1998 The mismeasure of memory: when retrieval fluency is misleading as a metamnemonic index. *J. Exp. Psychol. Gen.* **127**, 55–68. (doi:10.1037//0096-3445.127.1.55)

25 Dunlosky, J. & Metcalfe, J. 2010 *Metacognition.* Thousand Oaks, CA: Sage.

26 Finn, B. & Metcalfe, J. 2008 Judgments of learning are influenced by memory for past test. *J. Mem. Lang.* **58**, 19–34. (doi:10.1016/j.jml.2007.03.006)

27 Fleming, S. M. & Dolan, R. J. 2012 The neural basis of metacognitive ability. *Phil. Trans. R. Soc. B* **367**, 1338–1349. (doi:10.1098/rstb.2011.0417)

28 Kelly, C. M. & Jacoby, L. L. 2000 Recollection and familiarity: process-dissociation. In *The Oxford handbook of memory* (eds E. Tulving & F. I. M. Craik), pp. 215–222. London, UK: Oxford University Press.

29 Koriat, A. 2002 Metacognition research: an interim report. In *Applied metacognition* (eds T. J. Perfect & B. L. Schwartz), pp. 261–286. Cambridge, UK: Cambridge University Press.

30 Koriat, A. 2008 Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Mem. Cogn.* **36**, 416–428. (doi:10.3758/MC.36.2.416)

31 Miele, D. M., Wager, T. D., Mitchell, J. P. & Metcalfe, J. 2011 Dissociating neural correlates of action monitoring and metacognition of agency. *J. Cogn. Neurosci.* **23**, 3620–3636. (doi:10.1162/jocn_a_00052)

32 Cosentino, S., Metcalfe, J., Holmes, B., Steffener, J. & Stern, Y. 2011 Finding the self in metacognitive evaluations: a study of metamemory and agency in non-demented elders. *Neuropsychology* **25**, 602–612. (doi:10.1037/a0023972)

33 David, A. S., Bedford, N., Wiffen, B. & Gilleen, J. 2012 Failures of metacognition and lack of insight in neuropsychiatric disorders. *Phil. Trans. R. Soc. B* **367**, 1379–1390. (doi:10.1098/rstb.2012.0002)

34 Cosentino, S. A., Metcalfe, J., Butterfield, B. & Stern, Y. 2007 Objective metamemory testing captures awareness of deficit in Alzheimer's disease. *Cortex* **43**, 1004–1019. (doi:10.1016/S0010-9452(08)70697-X)

35 Fleming, S. M., Weil, R. S., Nagy, Z., Dolan, R. J. & Rees, G. 2010 Relating introspective accuracy to individual differences in brain structure. *Science* **329**, 1541–1543. (doi:10.1126/science.1191883)

36 Yokoyama, O. *et al.* 2010 Right frontopolar cortex activity correlates with reliability of retrospective rating of confidence in short-term recognition memory performance. *Neurosci. Res.* **68**, 199–206. doi:10.1016/j.neures.2010.07.2041

37 Blakemore, S.-J. 2003 Deluding the motor system. *Conscious. Cogn.* **12**, 647–655. (doi:10.1016/j.concog.2003.07.001)

38 Haggard, P. 2008 Human volition: towards a neuroscience of will. *Nat. Rev. Neurosci.* **9**, 934–946. (doi:10.1038/nrn2497)

39 Wolpert, D. M. 1997 Computational approaches to motor control. *Trends Cogn. Sci.* **1**, 209–216. (doi:10.1016/S1364-6613(97)01070-X)

40 Wolpert, D. M., Ghahramani, Z. & Jordan, M. I. 1995 An internal model for sensorimotor integration. *Science* **269**, 1880–1882. (doi:10.1126/science.7569931)

41 Blakemore, S.-J., Wolpert, D. M. & Frith, C. 2002 Abnormalities in the awareness of action. *Trends Cogn. Sci.* **6**, 237–242. (doi:10.1016/S13646613(02)01907-1)

42 Knoblich, J., Stottmeister, F. & Kircher, T. 2004 Self-monitoring in patients with schizophrenia. *Psychol. Med.* **34**, 1561–1569. (doi:10.1017/S0033291704002454)

43 Synofzik, M., Vosgerau, G. & Newen, A. 2008 Beyond the comparator model: a multifactorial two-step account of agency. *Conscious. Cogn.* **17**, 219–239. (doi:10.1016/j.concog.2007.03.010)

44 Bandura, A. 2001 Social cognitive theory: an agentic perspective. *Annu. Rev. Psychol.* **52**, 1–26. (doi:10.1111/1467-839X.00024)

45 Kirkpatrick, M., Metcalfe, J., Greene, M. & Hart, C. 2008 Effects of intranasal methamphetamine on metacognition of agency. *Psychopharmacology* **197**, 137–144. (doi:10.1007/s00213-007-1018-2)

46 Tricomi, E. M., Delgado, M. R. & Fiez, J. A. 2004 Modulation of caudate activity by action contingency. *Neuron* **41**, 281–292. (doi:10.1016/S0896-6273(03)00848-1)

47 Michotte, A. l963 *The perception of causality.* New York, NY: Basic Books.

48 Hubbard, T. L., Blessum, J. A. & Ruppel, S. E. 2001 Representational momentum and Michotte's 'launching effect' paradigm (1946/1963). *J. Exp. Psychol. Learn. Mem. Cogn.* **27**, 294–301. (doi:10.1037/0278-7393.27.1.294)

49 Blakemore, S. J., Frith, C. D. & Wolpert, D. M. 1999 Spatio-temporal prediction modulates the perception of self-produced stimuli. *J. Cogn. Neurosci.* **11**, 551–559. (doi:10.1162/089892999563607)

50 Blakemore, S. J., Wolpert, D. M. & Frith, C. D. 2000 Why can't you tickle yourself? *Neuroreport* **11**, R11–R16. (doi:10.1097/00001756-200008030-00002)

51 Schlottman, A. & Shanks, D. R. 1992 Evidence for a distinction between judge and perceived causality. *Q. J. Exp. Psychol.* **44**, 321–342. (doi:10.1080/02724989243000055)

52 Jenkins, A. C. & Mitchell, J. P. 2010 Medial prefrontal cortex subserves diverse forms of self-reflection. *Soc. Neurosci.* **6**, 211–218. (doi:10.1080/17470919.2010.507948)

53 Ochsner, K. N., Beer, J. S., Robertson, E. R., Cooper, J. C., Kihlstrom, J. F., D'Esposito, M. & Gabrieli, J. D. E. 2005 The neural correlates of direct and reflected self-knowledge. *Neuroimage* **28**, 797–814. (doi:10.1016/j.neuroimage.2005.06.069)

54 First, M. B., Spitzer, R. L., Gibbon, M. & Williams, J. B. W. 1995 *Structured clinical interview for DSM-IV axis I disorders, patient edition* (*SCID-P*), *version 2.* New York, NY: New York State Psychiatric Institute, Biometrics Research.