Metacognition and sense of agency

Wen Wen[1,2] , Lucie Charles[3], and Patrick Haggard[3]

[1]Department of Psychology, Rikkyo University, 1-2-26 Kitano, Niiza, Saitama 352-8558, Japan

[2]Department of Precision Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

[2]Institute of Cognitive Neuroscience, University College London, Alexandra House, 17 Queen Square, London WC1N 3AZ, UK

Corresponding author:

Wen Wen

Department of Psychology, Rikkyo University, 1-2-26 Kitano, Niiza, Saitama 352-8558, Japan; Email: wen@rikkyo.ac.jp

Abstract

Intelligent agents need to understand how they can change the world, and how they cannot change it, in order to make rational decisions for their forthcoming actions, and to adapt to their current environment. Previous research on the sense of agency, based largely on subjective ratings, failed to dissociate the sensitivity of sense of agency (i.e., the extent to which individual sense of agency tracks actual instrumental control over external events) from judgment criteria (i.e., the extent to which individuals self-attribute agency independent of their actual influence over external events). Furthermore, few studies have examined whether individuals have metacognitive access to the internal processes underlying the sense of agency. We developed a novel two-alternative-forced choice (2FAC) control detection task, in which participants identified which of two visual objects was more strongly controlled by their voluntary movement. The actual level of control over the target object was manipulated by adjusting the proportion of its motion that was driven by the participant's movement, compared to the proportion driven by a pre-recorded movement by another agent, using a staircase to hold 2AFC control detection accuracy at 70%. Participants identified which of the two visual objects they controlled, and also made a binary confidence judgment regarding their control detection judgement. We calculated a bias-free measure of first-order sensitivity (d') for detection control at any given level of participant's own movement. The proportion of pre-recorded movements determined by the staircase could then be used as an index of control detection ability. We identified two distinct processes underlying first-order detection of control: one based on instantaneous sensory predictions for the current movement, and one based on detection of a regular motor-visual relation across a series of movements. Further, we found large

individual differences across 40 particpants in metacognitive sensitivity (meta-d') even though first-order sensitivity of control detection was well controlled. Using structural equation modelling (SEM), we showed that metacognition was negatively correlated with the predictive process component of detection of control. This result is inconsistent with previous hypotheses that detection of control relies on metacognitive monitoring of a predictive circuit. Instead, it suggests that predictive mechanisms that compute sense of agency may operate unconsciously.
Keywords: sense of agency, metacognition, signal detection theory, confidence, control detection

1 **Introduction**

The subjective feeling of controlling one’s own actions, and, through them, events in the outside world, is called the sense of agency. It is a key feature of consciousness in humans, and perhaps also in some other animals. The sense of agency can be divided into two broad categories: Embodied sense of agency and external sense of agency (Grünbaum & Christensen, 2020; Schram & Grünbaum, 2018; Wen, 2019). Embodied sense of agency refers to the feeling of controlling our own body, movements, and sensations. It involves the integration of sensory information, such as proprioception, tactile, and kinesthetic cues, to create a coherent sense of the body as a self-contained entity. Embodied sense of agency is an essential part of body consciousness, which encompasses the awareness of controlling one's own body and its position in space. It is established during the long-term motor development and is usually robust. On the other hand, external sense of agency refers to the feeling of controlling events outside of the body, such as objects, other people, and the environment. It involves the attribution of causality and intentionality over external events, which allows us to make sense of our surroundings and interact with them in a meaningful way. External sense of agency plays an important role in our interaction with the external world. It allows us to predict the consequences of our actions, plan and execute complex tasks, and communicate with others effectively. It is also closely related to our sense of social identity and the way we perceive ourselves in relation to others. In this paper, we focus on the external sense of agency for perceptual and metacognitive judgments.

When we reach and grab a bottle of water, we know that we have located and then moved the water bottle. After pressing a button on a controller, the air conditioner came on, and we do not then doubt that we turned it on. We ubiquitously attribute external events to ourselves and to others based on our/their actions, and the sensory inputs following these actions. Action-outcome attribution can involve a low-level feeling, which may automatically influence our perception of sensory input, for example through sensory attenuation, or distortions of time perception (Blakemore et al., 1998; Haggard et al., 2002). Both views treat sense of agency as an association between a primary motor and a primary sensory event. In contrast, other studies suggested the sense of agency is *metacognitive*, meaning that it involves monitoring and detecting causal relationships between one's actions and events that are caused by those actions (Carruthers, 2012; Deroy et al., 2016; Metcalfe & Greene, 2007). More recent studies using formal definitions of metacognition based on perceptual confidence however failed to find evidence that agency judgements involve meta-level information, over and above primary sensorimotor information (Chambon et al., 2014; Constant et al., 2022).

The ability to monitor one's own cognitive processes is called metacognition (Metcalfe & Shimamura, 1995; Nelson, 1984). Perceptual metacognition concerns the ability that people have to evaluate the accuracy of that judgment of their own first-order perceptual decisions, by making a second-order decision about the likeliness that the first-order decision is correct. This is typically expressed by indicating the degree of confidence in the first-order decision. Metacognition is fundamental for humans to adapt to changes in the environment and to update their internal representations of the world (Nelson & Dunlosky, 1991). While individuals' metacognitive ability to make perceptual decisions has been studied extensively (Ais et al., 2016; Deroy et al., 2016; Fleming et al., 2012; Gardelle et al., 2016), little is known about the metacognition of sense of agency.

In particular, metacognitive sensitivity is intrinsically affected by the accuracy of first-order judgment (Fleming & Lau, 2014). Because the judgment of agency is an integration of multiple processes including both high and low level signals, and both internal and external cues (Moore & Fletcher, 2012; Wen et al., 2015), experimental manipulation of sense of agency remains difficult. Most previous studies asked participants to judge sense of agency using either rating scales (e.g., Ebert & Wegner, 2010; Sato & Yasuda, 2005; Wegner et al., 2004) or yes-no responses (e.g., Farrer et al., 2003; Maeda et al., 2012; Saito et al., 2017). Many previous studies measured how sense of agency was affected by experimenter-induced disturbances to the consequences of one's own action. These manipulations of self-caused events were designed to make the signals for self-control ambiguous. In such ambiguous settings, judgements of agency may reflect the dual influences of a participant's sensitivity to whether they actually have agentic control over the outcome, and the degree of their trait bias to self-attribute agency, irrespective of their actual degree of control (Wen, 2019; Wen & Imamizu, 2022).

In the present study, we measured the perceptual sensitivity of the sense of agency, as distinct from bias in agency judgments, using a forced-choice paradigm. We also measured the metacognition of sense of agency. In essence, we considered detection of one's control over an environmental object as a first-order judgment of agency, akin to classical perceptual judgement. We then considered participants' confidence in their agency detection as a second-order, metacognitive judgment, akin to classical perceptual confidence judgements (Fleming, 2017; Fleming et al., 2010; Fleming & Lau, 2014; Maniscalco & Lau, 2012). Through this approach, we could investigate how much the processes that involve detection of control are accessible to metacognition.

According to the influential comparator model of motor control, efference copies of motor commands generate a sensory prediction in the brain and is compared with the actual sensory input (Blakemore et al., 1998, 2002). This requires an internal model of the relation between motor commands and resulting sensory events. Sense of agency is reduced when the sensory input does not match the prediction: what happened is not what was predicted on the basis of the efference copy (Blakemore et al., 1998, 2002). Furthermore, previous studies showed that detecting a decrease in control is more salient than detecting an increase in control (Wen et al., 2021; Wen & Haggard, 2018). This indicates that the process of detecting control may differ from the process from detecting a loss of control (Wen et al., 2021). People can detect their control via a predictive process, in which they form a sensory prediction and continuously search for a match to the sensory prediction. Such prediction-based processes recall comparator model, but may have very different neural mechnisms (Wen et al., 2019). Previous studies reported that self-generated events attract attention (Kumar et al., 2015; Salomon et al., 2013; Wen & Haggard, 2018), indicating the possibility that the search for control can be conducted without attentional monitoring. Furthermore, we recently suggested that a second mechanism may contribute to sense of agency.

People can also detect a regular relationship between their own actions and external events even when they cannot predict the sensory outcome of any individual action, due to lacking an internal model (Wen & Haggard, 2020). Specifically, Wen and Haggard (2020) showed that people can recognize the object that contained their real-time mouse movements among several objects, even on trials where the dot motions on the screen were spatially rotated by 90 degrees relative to the mouse movements, and, crucially, when participants were prevented from simply learning the 90 degree visuomotor transformation by random trial-to-trial variations in angular bias. Whereas participants readily learn a 90 degree transformation when it is consistent (Imamizu et al., 2000), randomly interleaving 90 degree and 0 degree rotations made dot motion unpredictable, and prevented model learning. Participants' detection of control in such conditions cannot be based on model predictions. Instead, detection of control must occur because participants notice a regular relation between actions and sensory input, occurring over a series of several actions. This visuomotor correlation allowed them to have the sense of agency even though they cannot perform model-based control. Such a process reflects an online recalibration of the sense of agency (Stern et al., 2022). It is more retrospective than predictive and requires access to short-term memory. It is likely that conscious monitoring plays an important role in this regularity detection process.

In summary, the present study proposed a novel approach to measure both the sensitivity of control detection and the metacognition of control detection. We used a design that allowed us to distinguish the respective contributions to sense of agency of a predictive process and a regularity-based process. We could then ask which of these two processes supported higher metacognitive sensitivity, which implies gerater accessibility of high-level conscious thought. Participants continuously moved the computer mouse for 3 s, while they observed the movement of two visual objects on the computer screen. They were asked: over which object did you have more control? One of the objects (the distractor) moved on a pre-recorded trajectory entirely unrelated to the participant's actions. The other object (the target) moved according to a blend of the participant's own actions and pre-recorded movements. The proportions of the participant's action and of the pre-recorded movement could be controlled by the experimenter. Notice that all objects moved with the same velocity as the mouse movements, meaning that temporal cues are not useful for detecting which target was controlled. Following each control detection judgment, participants gave a yes/no metacognitive judgment regarding whether they were confident of their control detection judgement. A correct detection response is when the target was selected. This paradigm allows us to calculate the first-order d' using the signal detection theory (Green & Swets, 1966) (i.e., the discrimination sensitivity between a controllable object and an uncontrollable object) besides the judgment bias. In all the trials, we combined participants' mouse movements with a pre-recorded movement, the proportion of which was controlled by a staircase procedure. In half of the trials at random, we also added an angular bias of 90 degrees to the motion of the target. As a result, the displayed moving direction always deviated from the intended moving direction, creating significant prediction errors in the predictive process. However, this 90-degree rotation did not affect the spatial similarity of the spatial patterns between one's mouse movements and the displayed movements. For example, if someone draws a straight line with the mouse, a 90-degree rotation will result in a straight trajectory on the screen that is rotated by 90 degrees (assuming that no pre-recorded motion has been mixed in with one's own movement). In such cases, there is a regular relationship, or spatial correlation, between one's mouse movements and the displayed target trajectory (Wen & Haggard, 2020). On the other hand, consider a participant who makes very wiggly, oscillatory mouse movements. If they see that dot A moves in an oscillatory fashion, whereas dot B moves in a smoother way, they may judge that they control dot A. Crucially, they can use this heuristic both in a 0 degree angular bias condition, and in a 90 degree angular bias condition, and they can use it even if the angular bias varies randomly across trials. Thus, even in the absence of any predictive model relating the spatial features of their movement to the spatial displacements of the dot, they can use the regular relation between movement events, such as moments of high and low acceleration in an oscillating trajectory, and visual events, to identify which dot they control. We refer to this as the regularity process for the detection of agency. The ability to use regularity as a cue to agency presumably varies among individuals. Those who can better detect regularity are expected to perform better than those who are less proficient at detecting regularity, particularly when a varying angular bias prevents use of a learned predictive model.

In addition, confidence ratings allow us to calculate the metacognitive sensitivity (meta-d') of the sense of agency (Charles et al., 2013; Fleming & Lau, 2014; Rouault et al., 2018; Rounis et al., 2010). Importantly, we maintained the accuracy of the control detection at 70% by increasing/decreasing the proportion of pre-recorded motion in 5% steps, using a 2-up/1-down staircase (Fleming et al., 2010; García-Pérez, 1998). We could thus investigate metacognition of control while the first-level sensitivity (i.e., the accuracy of control detection) was held constant.

2 Methods

2.1 Participants

The study has been approved by the local ethics committee of the University of Tokyo. Participants were recruited by advertising in the university. Fifty-five participants took part in this experiment (22 females, average age = 23.9 year, SD = 7.0 years). All participants provided written informed consent before participation, and received reimbursement for their time.

Our pilot testing using a small number of repetitions (20 repetitions in each condition) confirmed a large effect size in the differences in the proportion of added pre-recorded movements and motor control performance between the two designed experimental conditions with 0° and 90° angular bias (Cohen's d = 0.78 and 1.24, respectively). However, due to the small number of trials, we were unable to acquire stable metacognitive measures in the pilot testing. To conduct correlation analyses between the first-order performance and metacognition among participants, we chose a sample size of 40 to ensure sufficient power.

A criterion of data exclusion was set before the experiment. Participants whose staircase hit the bottom (i.e. 0% of pre-recorded movement) in any of the conditions in the control detection task should be excluded. This case usually reflects poor control detection performance with an angular bias of 90°. In our pilot testing, three participants out of 20 matched this criterion for data exclusion in the condition of 90° angular bias. In the actual experiment, 15 participants out of 55 matched this criterion, resulting in a final sample size of 40.

2.2 Experimental Task and Procedure

Participants were tested individually, using a 27-in LCD monitor with a resolution of 1,920 pixels × 1,080 pixels (597 mm × 336 mm, width × height), a keyboard, and a computer mouse. They were seated approximately 60 cm from the monitor. Participants conducted two experimental tasks. The control detection task was the main task to measure the sense of agency and confidence and was conducted first. The reaching task was an additional task to measure the extent of prediction errors reflected by motor control performance and was conducted after the control detection task.

In each trial of the control detection task (**Figure 1A**), participants were shown a black 12.4-mm (40 pixels) square and a black 12.4-mm dot with a gray background. The two shapes were presented at random positions in the central area of the screen, with a maximum distance of 77.5 mm (250 pixels) from the center, and a minimum distance of 62 mm (200 pixels) between them. Participants were asked to move the mouse freely to trigger the movement of the two objects. The velocity, onset, and offset of the shapes' motion corresponded to participants' mouse movement, while the directions might be different from the mouse movement (Wen et al., 2020). In the case of the distractor, at each moment during the mouse movement, the x- and y-axes of one of the two shapes were a replay of pre-recorded movements, with the moving distance normalized to match the mouse movement. A section was randomly chosen from 50,000 pre-recorded continuous movements for each trial. In this case, participants did not have any control over the shape's moving direction. In the case of the target, the moving direction of the stimulus was a combination of participants' mouse movements and another section of pre-recorded movements (**Figure 2**). In this case, participants had an intermediate level of actual control over the stimuli. Specifically, at each moment during the mouse movement, the x- and y-axes of the pre-recoded movement were normalized to match the mouse movement, and then mixed with the mouse movement with certain weightings depending on the proportion of pre-recorded movements at each axis. For example, in the case of 70% pre-recorded movements, pre-recorded movements was mixed with participants' mouse movements at a 70/30 ratio. This method allows free manipulation of the proportion of self-motion between 0-100% (see S1 for a demonstration video showing how the actual displayed movements looked). When combined with an angular bias, the mixed movement will be rotated clockwise to a certain degree. In addition, during the calculation of stimulus movement, the magnitude of mouse movement was reduced to 1/5 to prevent excessive movement stimuli. At last, the stimuli were restricted within the screen when they reach the border of the screen. In such case, the algorithm re-selected a new section of pre-recorded movements that moves away from the border.

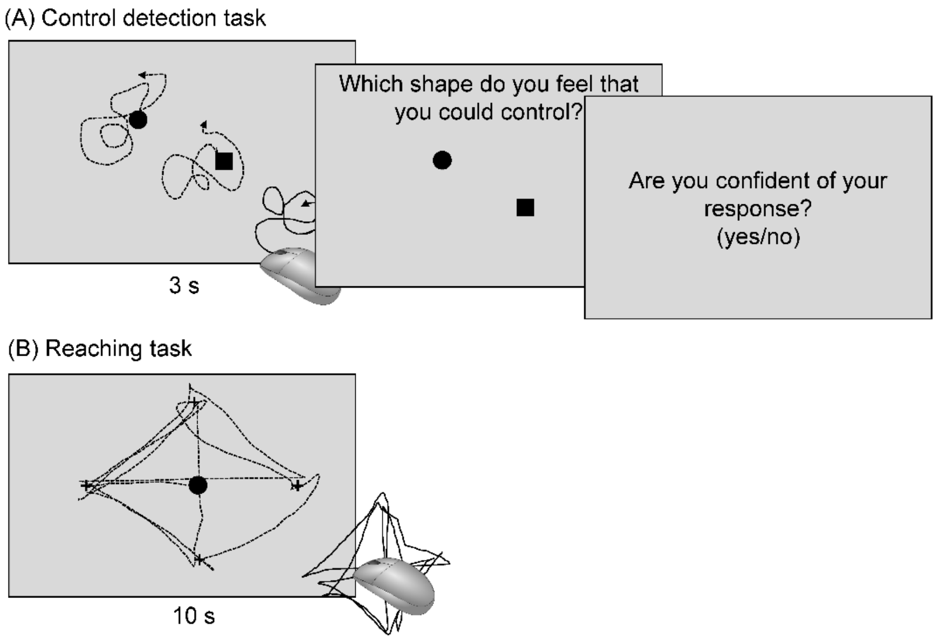

Figure 1. The timeline of the two experimental tasks.

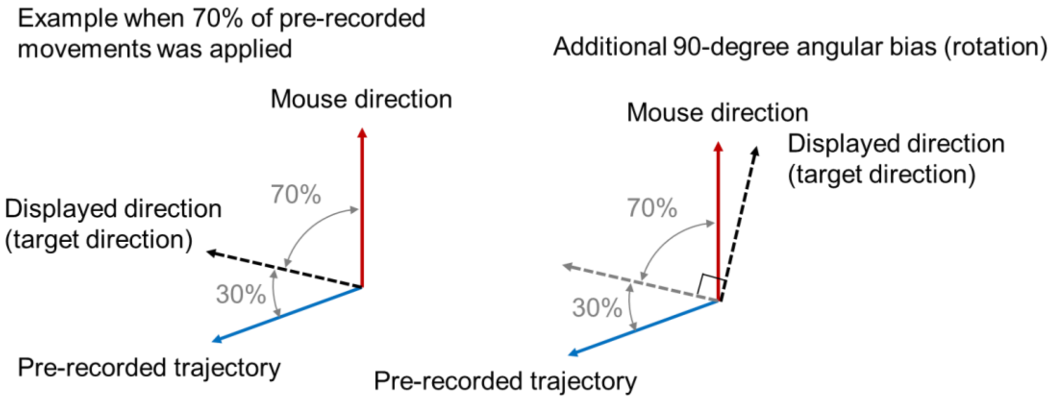

Figure 2. An example of one frame showing how the visual cursor might be generated. In this example, 70% of pre-recorded movements is applied (left). An additional 90-degree angular bias could also be applied (right).

In half of the trials, participants were controlling a square, the dot shape corresponding to the distractor, and in the other half of the trials, they were controlling a dot and the square was the distractor. Participants moved the two stimuli using the mouse for 3 s. Thereafter, they identified whether they felt they had more control over the square or over the dot, by pressing one of two labeled response keys on the keyboard. After this control detection response, they made a binary confidence judgment (low confidence vs high confidence) by pressing one of two labeled keys on the keyboard. The message of "next trial" on screen was shown 0.5 s after the confidence judgment.

In the control detection task, there were two conditions, one condition with 0° angular bias combined with a proportion of X% pre-recorded movements, and another condition with 90° angular bias combined with a proportion of Y% pre-recorded movements. X% and Y% were adjusted using a 2-up/1-down staircase procedure at 5% step, to hold the detection accuracy around 70%. Specifically, after each trial, the proportion of pre-recorded movements was increased by 5% if the following three conditions are all matched:

(1) participants had made two or more than two continuous correct responses;

(2) the average detection accuracy in the most recent 10 trials was above 70%;

(3) the proportion of pre-recorded movements was not increased in the last trial.

On the other hand, the proportion of pre-recorded movements was decreased by 5% if the following two conditions are all matched:

(1) participants made an incorrect response;

(2) the average detection accuracy in the most recent 10 trials was below 70%.

This staircase procedure has been confirmed to be effective to hold the performance comparable between the two conditions in our pilot testing. The initial value of X% and Y% was set to 60% and 45%, which was most likely to result in average detection accuracy of 70% across individuals according to our pilot results. Each condition was repeated 110 times. The initial 10 trials of each condition were excluded from analyses because the staircase was less stable during the initial stage. The proportion of pre-recorded movements was not changed for the first 4 trials, to calculate the initial detection accuracy. In addition, 30 extra trials with an angular bias of 30° and a fixed proportion of 40% of pre-recorded movements were mixed, to prevent adapting to certain angular transformations. The total number of trials was 250 in the control detection task. The trial order was fully randomized for each participant, and the staircase was manipulated separately in the background for each condition. The trials were divided into three blocks, containing 80, 80, and 90 trials, respectively. Participants took a break between blocks. A large number of trials were to ensure a stable measure of metacognition. Participants practiced for 6 trials, containing 3 trials of each condition with the proportion of pre-recorded movements fixed at the initial value.

After the control detection task, participants performed the reaching task (**Figure 1B**), with two conditions that match the averaged value of the proportion of pre-recorded movement in the control detection task. The proportion of pre-recorded movements used in the reaching task was constant for each participant, but differed among individuals according to their performance in the control detection task. The reaching task provided a test of whether participants can adapt to the 90-degree angular bias under conditions where 0-degree and 90-degree angular biases are randomized. Successful adaptation would imply comparable reaching performance in the 0-degree and 90-degree conditions after repeated exposure, whereas little or no adaptation would imply consistently worse reaching performance in the 90-condition.

In each trial of the reaching task, a dot was shown at the center of the screen with a 5.3 mm (20 pixels) black cross. The actual control of this dot was exactly as one of the two conditions in the control detection task. A target cross was randomly presented at one of four positions, which are 125.6 mm (405 pixels) horizontally or vertically away from the center of the screen. Participant moved the dot to touch the cross as quickly as possible by moving the mouse. The cross disappeared once it is touched by the dot, and then re appeared at a new possible out of the rest three possible positions. Participants were instructed to move the dot to touch the cross as many times as possible in each trial. Each trial lasted for 10 s from the onset of the first mouse movement. Each condition was repeated 10 times, and 10 extra trials with an angular bias of 30° and 40% of pre-recorded movements were mixed. There were therefore a total of 30 reaching trials. Participants practiced for 6 trials, containing 2 trials of each angular bias condition (0-degree, 30-degree, and 90-degree), before the actual reaching task. The order of the trials was randomized. The whole experiment took 50-60 min for each participant.

2.3 Data analyses

Results from 200 trials (excluding the 20 initial trials and the 30 extra trials) in the two experimental conditions in the control detection task were used for the estimation of perceptual sensitivity (d') and bias in the first-order judgment using Signal Detection Theory (Green & Swets, 1966). This perceptual sensitivity reflects the ability to distinguish an object that is under partial control from an object that is under no control. Because the task is designed to hold the detection accuracy constant at around 70%, we expect no difference in the first-order judgment between the conditions of 0° and 90° angular bias.

Furthermore, by combining the response of control detection and the confidence rating, we can estimate the meta-d' as the index of metacognitive sensitivity using the fitting method (Charles et al., 2013; Fleming & Lau, 2014; Rouault et al., 2018; Rounis et al., 2010). The meta-d' approach is based on an ideal observer model of the link between type 1 and type 2 signal detection theory. Given a particular type 1 variance structure and bias, the method fits type 2 responses (i.e. confidence rating) into a family of type 2 ROC curves and determines the best fitting curve to find out meta-d' (Fleming & Lau, 2014). The meta-d' can be expressed relative to d' to provide a measure of metacognitive efficiency. We compared the metacognitive sensitivity between the two experimental conditions. In addition, to analyse metacognitive bias, we retrieved the confidence criteria (C2_S1 and C2_S2), corresponding to the boundary the two confidence levels separating the low confidence responses from the high confidence responses for S1 (target is the noncontrollable object) and for S2 (target is the controllable object) respectively. We then computed the distance of the averaged confidence criteria for S1 and S2 to the first-order decision criterion (C1), which is denoted *C2*, to examine the distance between the confidence and first-order criterion (Charles et al., 2020; Sherman et al., 2018). In addition, because the confidence axis is not linear, we computed the log value of the distance as

$$\text{logC2} = (\log(-\text{C2\_S1}) + \log(\text{C2\_S2}))/2$$

The closer the confidence criterion is to the decision threshold, the more participants will have a tendency to report high confidence (Charles et al., 2020).

Next, to measure individual differences in the ability to use the prediction-based process and regularity-based process, we calculated the individual threshold of each process. The prediction-based process compares the intended motion (i.e., predicted moving direction) with the displayed motion. Therefore, the *angular error* between one's mouse movement and the displayed movement is useful for the predictive process. The larger the difference in angular error between the target and the distractor, the easier it is to detect the target using the predictive process. Figure 3A and 3B show the plots of angular errors against proportions of pre-recorded movements for the target in the condition of 0-degree angular bias and 90-degree angular bias, respectively. The angular error was useful for detecting the target only in the 0-degree condition. Because the staircase procedure ensured that the proportion of other's movement is adjusted so that accuracy is at the detection threshold, we can calculate the individual threshold of target detection based on the predictive process by simply averaging the value of angular errors at all the reversal points of the staircase in the condition of 0-degree angular bias for each individual (García-Pérez, 1998). This value was used as the index of *efficiency of predictive process* (a larger value refers to *higher* efficiency). **Figure 4A** shows the histogram of this measure of the *efficiency of predictive process* i.e. the angular errors for detecting the target at 70% accuracy in the condition of 0-degree angular bias.

In the case of 90-degree condition, angular errors are no longer useful to distinguish the target from the distractor. On the other hand, the spatial correlation between one's movements and the displayed movements can still be useful. Figure 3C and 3D show the plots of *bidimensional correlation* against proportions of pre-recorded movements for the target in each condition. Bidimensional correlation is often used to evaluate how similar two 2-D spatial patterns are (Tobler, 1965; Wen et al., 2013). Bidimensional correlation can therefore be used to measure the contribution of the regularity mechanism, in both 0-degree and 90-degree conditions. Bidimensional correlation requires longer samples than angular errors, because correlations can only be estimated reliably when many datapoints are available. Therefore, bidimensional correlation is most likely to be used in the 90-degree condition, when the more efficient, faster method of angular error detection is not available. In the condition of 90-degree angular bias, in contrast, adjusting the proportion of other's motion does not affect angular errors, since these are always at the maximal level, Figure 3B, but only affects the bidimensional correlation. Therefore, the staircase in the condition of 90-degree ensured that the input of bidimensional correlation achieved the threshold for detecting the target using the regularity detection process. We averaged the value of bidimensional correlations at reversal points in the condition of 90-degree angular bias for each individual and used this average value as the index of *efficiency of regularity* *detection* (larger value in this case refers to *poorer* efficiency: note the inverse relation compared to the predicted process). **Figure 4B** shows the histogram of bidimensional correlations for detecting the target at 70% accuracy in the condition of 90-degree angular bias. The individual variance when using bidimensional correlation (i.e., regularity cues) for control detection is relatively larger than when employing prediction errors (i.e., predictive cues) for the same task.

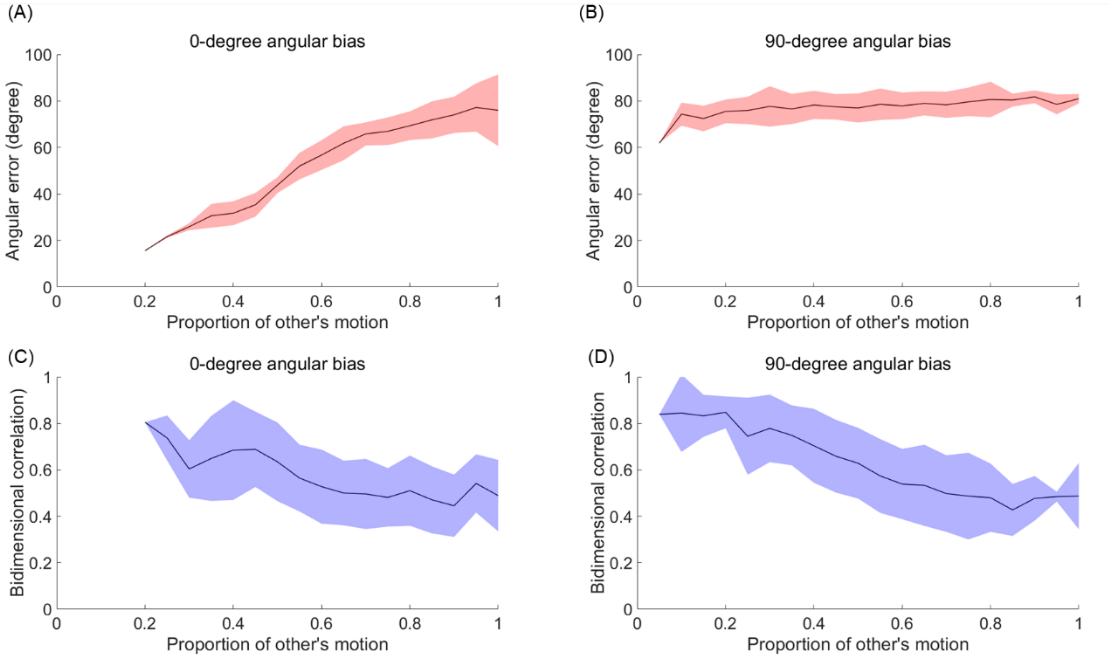

Figure 3. Plot of angular errors (A and B) and bidimensional correlations (C and D) against proportions of pre-recorded movements in each angular bias condition. Angular errors increase linearly when the proportion of pre-recorded movements increases in the 0-degree angular bias condition and remain constantly at their maximum in the 90-degree angular bias condition. On the other hand, bidimensional correlations decrease linearly when the proportion of pre-recorded movements increases in both angular bias conditions.

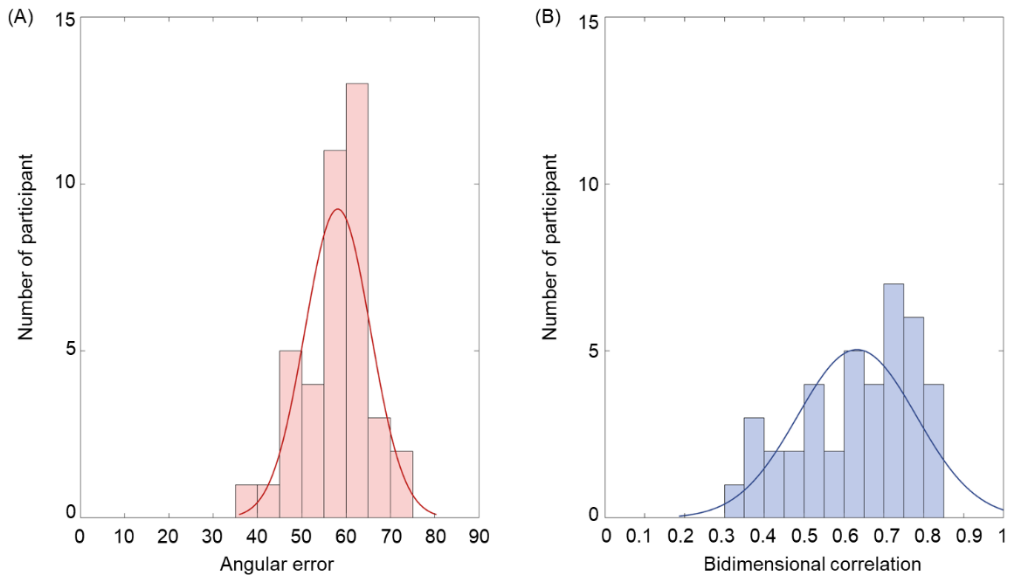

Figure 4. Histograms of angular errors (A) and bidimensional errors (B) for detecting the target dot with 70% accuracy in the conditions of 0-degree angular bias and 90-degree angular bias, respectively. Fitted normal distribution curves are also shown.

We next investigated whether individual differences in metacognition of agency were related to differencs in the efficiency of the predictive and regularity-based processes, using structure equation modelling (SEM) (see Results and **Figure 6** for the details for the model). The SEM contained four variables: the efficiency of the predictive process, the efficiency of regularity detection (see above), the sensitivity of sense of agency (d’) and the sensitivity of metacognition (meta-d’). The SEM was conducted to examine individual differences in how agency was computed, rather than the effect of manipulating experimental condition. Furthermore, as no significant differences in d’ or meta-d’ were found between the 0-degree angular bias and 90-degree angular bias conditions , d’ and meta-d’ were averaged across the two conditions for each participant. We focused on whether the efficiencies of the two internal processes for agency computation are associated with metacognitive awareness.

3 Results

Figure 5 shows the results of d' and meta-d', corresponding to the sensitivity of control detection judgement, and of metacognitive judgment of agency respectively. Mean d' was successfully controlled at a constant level in our task. A 2×2 (type of sensitivity, d' vs meta-d' × angular bias, 0 vs 90) repeated-measures ANOVA revealed neither significant main effect nor significant interaction (for the main effect of type of sensitivity: $F(1, 39) = 1.55$, $p = .220$, partial $\eta^2 = 0.038$; for the main effect of angular bias, $F(1, 39) = 0.22$, $p = .638$, partial $\eta^2 = 0.006$; for the interaction: $F(1, 39) = 0.011$, $p = .918$, partial $\eta^2 < .001$). The sensitivity of control detection and metacognition of control was thus comparable. This means on average the metacognitive sensitivity did not differ between the conditions when people detect their control using prediction-based processes and regularities. Bayes factor estimations (using LearnBayes package in R) showed that the posterior probability of no difference in d' and meta-d' between the two experimental conditions was high ($\tau = .01$; BF = .500 and .500, respectively). Furthermore, as shown in **Figure 5**, there were large individual differences in meta-d' while d' was well controlled. We suggest that a key source of individual differences may lie in the degree to which each individual can access the internal processes of prediction and regularity-detection that underlie control detection. In addition, the mean type 1 beta (criterion) was -.07 (*SD* = .19) and -.13 (*SD* = .19) in the 0-degree condition and the 90-degree condition, respectively. The average log*C2* was -0.20 (*SD* = .20) and -0.22 (*SD* = .25) for the 0-degree condition and the 90-degree condition, respectively. There was no significant difference in m-dist across the conditions ($t(39) = 1.004$, $p = .322$, Cohen's d = .12).

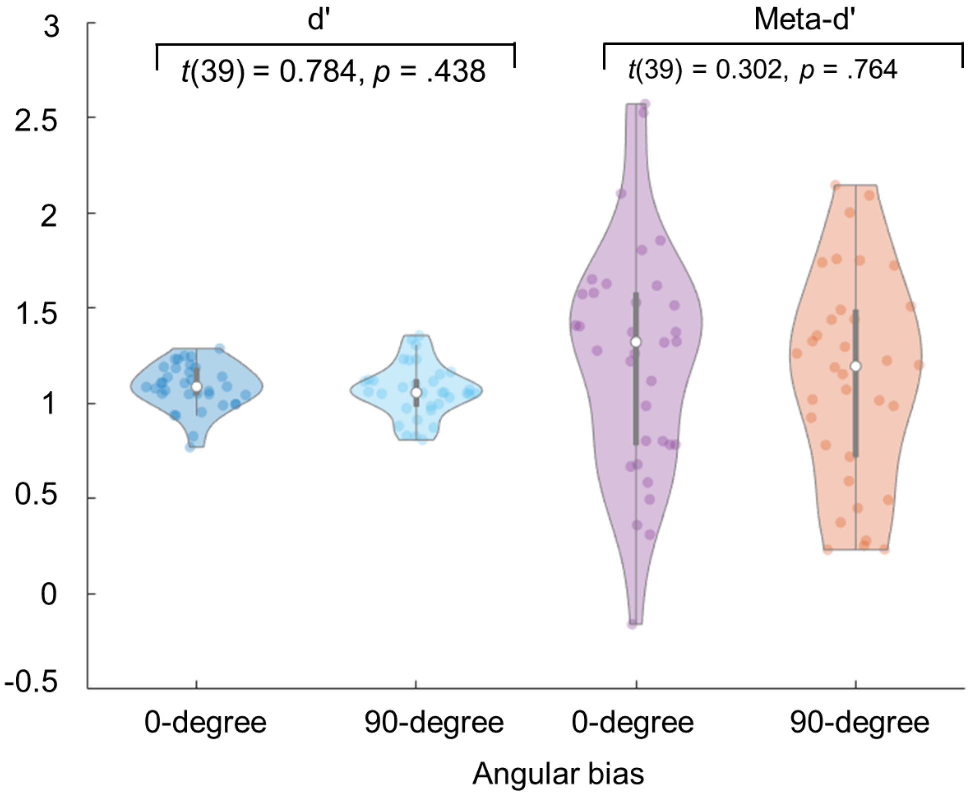

Figure 5. The results of d' and meta-d' in the two angular bias conditions.

We considered how sensitive each individual's sense of agency was to the true degree of control over the visual objects. As we argued above, this could depend on two different cognitive processes: the predictive process, and the regularity detection process. Further, these two processes could be associated with different levels of metacognitive awareness. To examine this hypothesis, we entered d', meta-d' and our estimates of the efficiency of each of the component internal processes into a structural equation model (**Figure 6A**). Note that in the predictive process, larger values indicate greater efficiency. In contrast, in the regularity detection process, larger values indicate less efficiency in detecting regularities. The standardized coefficient of each path within the SEM, the plot of individual efficiency of the predictive process against meta-d', and the plot of individual efficiency of regularity detection against meta-d' are shown in **Figure 6A**, **B**, and **C**, respectively. Because the control detection accuracy was controlled using the staircase, there is relatively little variance in d', we did not therefore focus on the paths involving d', although a path from d' to meta-d' was included on the theoretical grounds that metacognitive judgement is based on monitoring first-level information (Fleming & Lau, 2014). The model-fit criteria and fit indices are shown in **Table 1**.

First, as predicted, we found that efficiency of the predictive process significantly and *negatively* influenced metacognitive sensitivity (meta-d') (β = -.358, p = .016; r = -.331, p = .037). That means, when first-level detection accuracy was well controlled among individuals, people who were more efficient in detecting control using prediction-based process tended to show less accurate confidence judgments, indicating that they have poorer metacognitive access to the predictive processes. In other words, people who are highly sensitive to predictive cues to control may not closely track their internal processes of detecting control in their metagocognitive awareness.

On the other hand, efficiency of regularity detection also varied widely across people, but was less related to meta-d' (β = .172, p = .249; r = .058, p = .722). To directly compare the path linking the predictive process to meta-d' with the path linking the regularity mechanism to meta-d', we constrained the coefficients of the two paths to be equal in one model (i.e., the *constrained* model) and compared the model-fit indices with the proposed, unconstrained model in which metacognition could be differentially related to the predictive and regularity-based processes (i.e., the unconstrained model).

Only the proposed (uncontrained) model showed good fit to the data, indicating that the strength of the two paths are significantly different (**Table 1**). Regularity is another perceptual cue for people to detect their control besides prediction errors, especially in the condition of 90-degree angular bias when the angular error is always large but the spatial pattern can still be recognized. Individual differences in the efficiency of regularity detection did not affect metacognitive sensitivity. Thus, we found evidence that strong reliance on predictive processes for computing sense of control was associated with lower metacognitive awareness, but we found no evidence for any particular relationship between regularity detection processes and metacognition..

Table 1. Model-fit criteria and the fit indices of the proposed model and the constrained model. The proposed model closely fits the data.

	χ^2/df	GFI	AGFI	RMSEA
Criterion of good fit	< 2.0	> .90	> .90	< .08
Unconstrained model	1.178	.972	.858	.067
Constrained model	65.185	.654	.136	1.283

Note: The unconstrained model was the proposed model. The constrained model was used to compare the path linking the predictive process to meta-d' with the path linking the regularity mechanism to meta-d'.

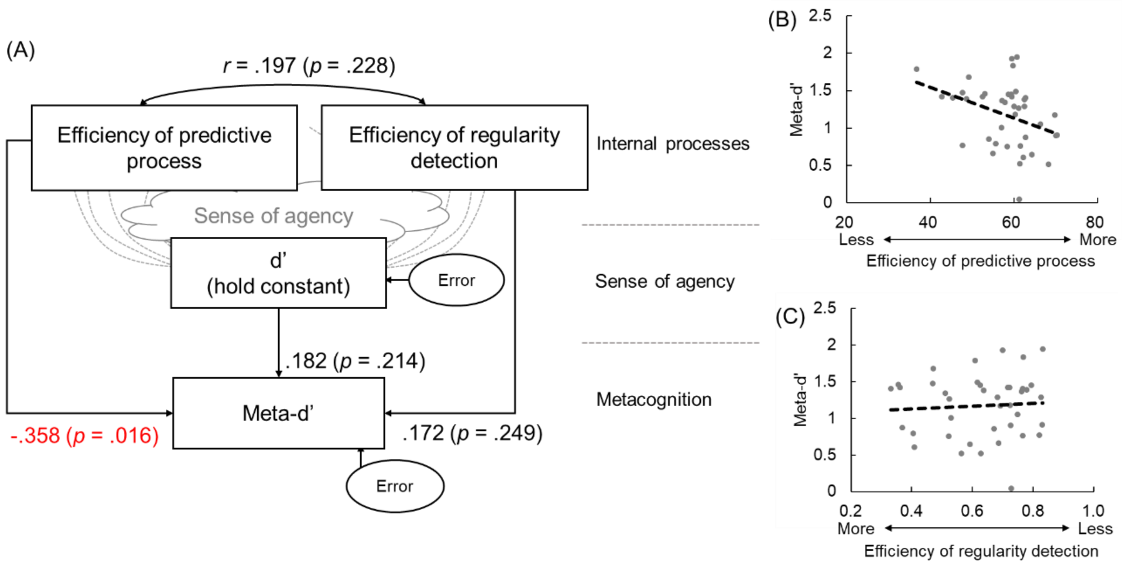

Figure 6. Structural equation model (SEM) of metacognitive sensitivity and the efficiencies of two internal processes (A), the plot of efficiency of predictive process against meta-d' (B), and the plot of efficiency of regularity detection against meta-d' (C)

Finally, we considered the possibility that participants may have adapted to the 90-degree angular bias after repeated exposure. In that case, they would not need to switch to regularity detection in the 90-degree condition but could use prediction-based processes from an updated internal model as cues for control detection. We planned that motor adaptation should not occur, by randomly switching between the two angular biases. However, we verified whether this manipulation worked by analyzing a reaching task performed after the main control detection task. If participants could learn a new model for the 90-degree angular bias, then their reaching performance would gradually become as accurate in the 90-degree as in the 0-degree condition. In fact, the gap in task performance between the two conditions did not reduce with experience (**Figure 7,** the grey line). A repeated-measures ANOVA of on the difference in task performance between the two conditions revealed no significant differences between the

10 successive presentations of the reaching task ($F(9, 351) = 0.885$, $p = .539$, partial $\eta^2 = 0.022$). This suggests that randomization of angular biases successively prevented motor adaptation.

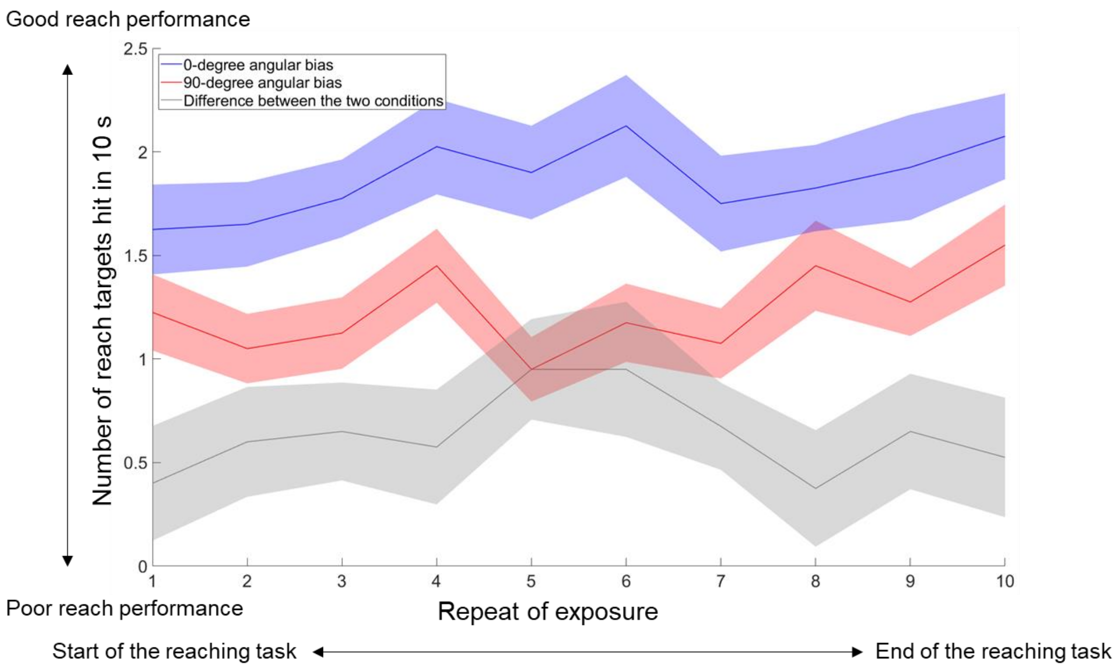

Figure 7. The number of successful hits on target in a reaching task presented at different stages of the reaching task. Note that participants do not adapt to the 90-degree angular bias, and their reaching does not improve even when they were repeatly exposed to the angular bias. Transperate background represents standard errors between participants.

4 Discussion

The present study aimed to dissociate sensitivity in sense of agency from judgment bias based on signal detection theory, and to compare sensitivity in first-order control detection with the metacognitive sensitivity of agency judgement (i.e., sensitivity of confidence judgements about one's own agency).. We developed a novel paradigm in which participants identified which of two visual objects they more strongly controlled. This paradigm allowed us to calculate the perceptual sensitivity (d') for identifying one's own agency based on signal detection theory (Green & Swets, 1966). By randomly interleaving trials with and without a 90-degree visuomotor bias transformation, we could distinguish two distinct cognitive mechanisms contributing to sense of agency, one based on prediction of the visual consequences of the current movement, and one based on regularities between multiple movements and their corresponding visual consequences. Specifically, a prediction-based method cannot correctly detect agency when the visomotor transformation is volatile, yet a regularity-based method remains viable. Furthermore, after each detection, we asked participants to give a confidence rating of their response, which we used to calculate the metacognitive sensitivity (meta-d') of the sense of agency, again using a signal detection theoretic approach (Maniscalco & Lau, 2012).

There are several interesting and novel findings from our results. First, we found large individual differences in how participants detected their own agency, with some participants relying more on prediction-based processes, and others relying more on regularities defined across multiple movements. Both mechanisms showed large individual differences in sensitivities. Second, metacognitive sensitivities showed comparable mean levels to first-order control detection sensitivities, showing that the internal processes of sense of agency are highly available for conscious access. Third, there were large individual differences in metacognitive sensitivity even when first-order sensitivity was well controlled among individuals. These individual differences in metacognition were significantly and negatively correlated with the individual differences in the effiency of the predictive component underlying the sense of agency, but were not correlated with the efficiency of regularity detection. These results showed that, while predictive mechanisms can provide sensitive first-order agency detection performance, they do not support accurate computation of confidence: participants who were excellent at control detection with very noisy sensory input using the prediction-based process in fact showed *poorer* confidence judgements than participants with less good control detection. In contrast, individual differences in the efficiency of regularity detection were also substantial, but did not significantly correlate with metacognitive sensitivity. This pattern of results shows that the regularity-based process is more consistently accessible to metacognition across individuals than the prediction-based process. People who are good at regularity detection do not show better or poorer megacognitive sensitivies than people who can not efficiently use regularities. We next extend our discussion on these findings.

First, regarding the individual differences in the sense of agency, results of previous studies using explicit agency ratings were likely to be affected by individual criteria of judgment (see Wen & Imamizu, 2022 for a review). For example, most studies show a general bias towards self-attribution of agency (Tsakiris et al., 2005). Individuals may express this general bias to varying extents. Because of such biases, indirect measures such as sensory attenuation and intentional binding have been proposed as alternative measures of sense of agency. However, these measures depend on perceptual domains such as tactile sensation or time perception. It remains unclear whether individual differences in perceptual sensitivity to these underlying domains affects sense of agency. Some recent studies using a control detection paradigm have used standard frameworks for perceptual sensitivity and metacognition to study sense of agency (Constant et al., 2022; Wang et al., 2021; Wen & Haggard, 2020; Wen & Imamizu, 2022). The present study further provided a method for studying metacognition of sense of agency while controlling the first-order sensitivity of the sense of agency.

Our experimental design also allowed us to identify specific cognitive mechanisms underlying the sense of agency, based on sensory predictions and on regularity detection. We found that individual differences in the efficiency of the regularity detection mechanism were strikingly large. Fifteen participants out of 55 originally tested were excluded from data analyses because they failed to detect the object they controlled in the 90-degree angular bias condition even when the proportion of pre-recorded movements approached 0%. The excluded participants were presumably unable to use the regularity detection process to identify the target, even when the difference in regularity between the target and the distractor was maximal.

Secondly, we found that the metacognitive sensitivity was at a comparable level as the first-order detection sensitivity on average. This may seem surprising. According to computational models of confidence judgment (Kepecs & Mainen, 2012; Meyniel et al., 2015; Pouget et al., 2016; Sanders et al., 2016), metacognitive judgments depend on reviewing the reliability of the first-order evidence, and estimating the likelihood of a correct or incorrect decision based on that evidence. The metacognitive accuracy was 63.2% (*SD* = 8.5%, range = 41 – 81%) and 62.5% (*SD* = 9.9%, range = 41 – 78%) in the condition of 0- and 90-degree angular bias, respectively, which was only slightly lower than the first-order detection accuracy (for 0-degree angular bias, *M* = 70.4%, *SD* = 2.0%, range = 65-75%; for 90-degree angular bias, *M* = 70.0%, *SD* = 2.4%, range = 63-75%). However, the individual difference in metacognitive sensitivity was large considering the fact that the first-order detection sensitivity was well controlled among individuals. In addition, we used angular bias manipulations to encourage participants to use regularity, rather than prediction, for control detection. These manipulations did not result in significant mean differences in meta-d’. This was not predicted and seems inconsistent with our findings that the predictive processes are less accessible to metacognition. Some participants showed higher metacognitive sensitivity than first-order sensitivity. This implies that, in error trials, they were able to detect they made a mistake and indicated that they were not confident in their first-order decision. This is usually the case when time-pressure is applied to the first decision – yet no explicit time pressure was present in our task. If we assume that accidentally pressing the wrong response key was rare, we might ask: if people knew they were wrong without being given more evidence, why didn’t they make the correct response from the beginning? Some participants might still have adopted spontaneously a strategy where they made fast guesses that they then evaluated as incorrect when reporting their confidence.

In addition, the number of trials in our study is relatively smaller compared to previous studies (e.g., 210 trials in Rouault et al., 2018; 300 trials in Rounis et al., 2010), and the number of trials may affect the test-retest reliability (Guggenmos, 2021). The relatively small number of trials in our study is due to the longer duration of each trial. Despite this, the estimation of meta-d' in our study still effectively captures the distribution of individual metacognitive sensitivity.

Given these substantial individual differences in agency sensitivity, we could examine the likely sources of metacognition of agency. In particular, we asked whether the confidence in agency judgements (captured by the meta-d’ measure) was related to a prediction-based process or a regularity-based process. We reasoned that low metacognitive sensitivity could arise if the processes of control detection were not tracked by metacognition – perhaps because they occurred below the level of conscious metacognitive monitoring. The results of SEM confirmed this possibility: We found a significant and negative influence of individual efficiency of the predictive process on meta-d'. This supports our hypothesis that the internal processes of control detection may not be accessible to metacognition. Prior research on the metacognition of visual detection also reported large individual differences in metacognitive sensitivity while the first-order detection accuracy was held constant (Fleming et al., 2010). Previous literature on sense of agency has suggested a strong link between conscious experience and predicting the outcomes of action (Blakemore et al., 1998). Our findings suggest that these links apply to first-level cognition, but may not apply to metacognition.

Because our stimuli were adapted to each individual, the evidence regarding control would be lower for more sensitive individuals. Could this lead to a trivial negative correlation between first-order sensitivity and metacognitive sensitivity? Metacognition relies on access to internal processes during the first-level judgment. If the internal processes for the first-level judgment are fully accessible to metacognition, then metacognitive sensitivity should be comparable to the first-level performance, regardless of how challenging the first-order judgment is. The lack of a significant correlation between the sensitivity of regularity detection and metacognitive sensitivity indicates that some people can detect their control under very difficult conditions, even when the regularity (measured by the bidimensional correlation) is low. In such cases, their metacognitive sensitivity is not reliably lower than that of people who require high regularity to detect control. This means that the internal processes of regularity detection are probably equally accessible by metacognition among individuals regardless of their efficiencies. Regularity detection requires the ability to establish hypotheses regarding the relationship between one's own motion and the sensory feedback, and actively sample evidence to test these hypotheses across multiple actions. Consciousness monitoring is probably required for these processes.

Previous research has discussed the link between the sense of agency and metacognition (Carruthers, 2012; Constant et al., 2022; Krugwasser et al., 2022; Metcalfe et al., 2013; Metcalfe & Greene, 2007). Earlier studies suggested that the sense of agency is elicited by the internal monitoring of intentions and outcomes (Metcalfe et al., 2013), and multiple cues are integrated during the judgment of agency (Moore & Fletcher, 2012; Synofzik et al., 2009). However, no study has ever examined to which extent people can access the internal processes underlying the sense of agency. A recent study using an exploratory modelling approach showed that although the noise level at the sensory input influenced the explicit judgment of agency, in the same way that noise affects perceptual confidence ratings, the judgment criterion of agency does not involve metacognitive noise estimates (Constant et al., 2022). In other words, agency judgments reflect the first-order process of the sensory input rather than metacognitive monitoring of the internal processes. The present study further showed that the first-order internal processes underlying the sense of agency may not be fully accessible by metacognition. Particularly in very sensitive people, the first-order detection of control can occur without metacognition.

In conclusion, the present study is the first to examine how metacognition of control is linked to individual sensitivity of the sense of agency. We found large individual differences in the sensitivity of sense of agency, in both the processes of detecting control via sensory predictions and regularities. Moreover, we found that people who have highly efficient predictive processes probably detect control using a circuit that does not support metacognitive access. On the other hand, the detection of regularity is probably more accessible by metacognition. Detecting one's control in the environment plays an important role in motor learning, decision making, and social interaction. Research on the developmental processes and abnormalities of sense of agency and metacognition are expected to understand why an accurate sense of agency is important for humans.

Acknowledgements

This work was supported by JSPS KAKENHI (Grant Number 19K20642 and 21H03780) to WW, a British Academy Post-doctoral Fellowship to LC, and an ESRC grant ES/V00378X/1 to LC and PH.

Declaration of interest

The authors declare that there is no conflict of interest.

Supplementary material

S1. A demonstration video of the stimuli movement generated from one of the actual trials. Note that the picture of computer mouse in the video was not presented in the actual task. The angular bias was 0 degree and the proportion of pre-recorded motion was 65% in the video.

References

Ais, J., Zylberberg, A., Barttfeld, P., & Sigman, M. (2016). Individual consistency in the accuracy and distribution of confidence judgments. *Cognition*, *146*, 377–386. https://doi.org/10.1016/j.cognition.2015.10.006

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*(7), 635–640. https://doi.org/doi:10.1038/2870

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (2002). Abnormalities in the awareness of action. *Trends in Cognitive Sciences*, *6*(6), 237–242. https://doi.org/10.1016/S1364-6613(02)01907-1

Carruthers, G. (2012). A metacognitive model of the sense of agency over thoughts. *Cognitive Neuropsychiatry*, *17*(4), 291–314. https://doi.org/10.1080/13546805.2011.627275

Chambon, V., Filevich, E., & Haggard, P. (2014). What is the human sense of agency and is it metacognitive. In *The cognitive neuroscience of metacognition*. Springer. https://doi.org/10.1007/978-3-642-45190-4

Charles, L., Chardin, C., & Haggard, P. (2020). Evidence for metacognitive bias in perception of voluntary action. *Cognition*, *194*, 104041. https://doi.org/10.1016/j.cognition.2019.104041

Charles, L., Van Opstal, F., Marti, S., & Dehaene, S. (2013). Distinct brain mechanisms for conscious versus subliminal error detection. *NeuroImage*, *73*, 80–94. https://doi.org/10.1016/j.neuroimage.2013.01.054

Constant, M., Salomon, R., & Filevich, E. (2022). Judgments of agency are affected by sensory noise without recruiting metacognitive processing. *ELife*, *11*, 1–28.

https://doi.org/10.7554/eLife.72356

Deroy, O., Spence, C., & Noppeney, U. (2016). Metacognition in Multisensory Perception. *Trends in Cognitive Sciences*, *20*(10), 736–747. https://doi.org/10.1016/j.tics.2016.08.006

Ebert, J. P., & Wegner, D. M. (2010). Time warp: Authorship shapes the perceived timing of actions and events. *Consciousness and Cognition*, *19*(1), 481–489. https://doi.org/10.1016/j.concog.2009.10.002

Farrer, C., Franck, N., Paillard, J., & Jeannerod, M. (2003). The role of proprioception in action recognition. *Consciousness and Cognition*, *12*, 609–619. https://doi.org/10.1016/S1053-8100(03)00047-3

Fleming, S. M. (2017). HMeta-d : Hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings. *Neuroscience of Consciousness*, *March*, 1–14. https://doi.org/10.1093/nc/nix007

Fleming, S. M., Huijgen, J., & Dolan, R. J. (2012). Prefrontal contributions to metacognition in perceptual decision making. *Journal of Neuroscience*, *32*(18), 6117–6125. https://doi.org/10.1523/jneurosci.6489-11.2012

Fleming, S. M., & Lau, H. C. (2014). How to measure metacognition. *Frontiers in Human Neuroscience*, *8*(July), 1–9. https://doi.org/10.3389/fnhum.2014.00443

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*(5998), 1541–1543. https://doi.org/10.1126/science.1191883

García-Pérez, M. A. (1998). Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties. *Vision Research*, *38*(12), 1861–1881. https://doi.org/10.1016/S0042-6989(97)00340-4

Gardelle, V. De, Corre, F. Le, & Mamassian, P. (2016). Confidence as a common currency between vision and audition. *PLoS ONE*, *11*(1), e0147901. https://doi.org/10.1371/journal.pone.0147901

Green, D., & Swets, J. (1966). *Signal detection theory and psychophysics*. NY: Wiley.

Grünbaum, T., & Christensen, M. S. (2020). Measures of agency. *Neuroscience of Consciousness*, *2020*(1), 1–13. https://doi.org/10.1093/NC/NIAA019

Guggenmos, M. (2021). Measuring metacognitive performance: type 1 performance dependence and test-retest reliability. *Neuroscience of Consciousness*, *2021*(1), 1–14. https://doi.org/10.1093/nc/niab040

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience*, *5*(4), 382–385. https://doi.org/10.1038/nn827

Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., Pütz, B., Yoshioka, T., & Kawato, M. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature*, *403*(6766), 192–195. https://doi.org/10.1038/35003194

Kepecs, A., & Mainen, Z. F. (2012). A computational framework for the study of confidence in humans and animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1594), 1322–1337. https://doi.org/10.1098/rstb.2012.0037

Krugwasser, A. R., Stern, Y., Faivre, N., Harel, E. V., & Salomon, R. (2022). Impaired sense of agency and associated confidence in psychosis. *Schizophrenia*, *8*, 32. https://doi.org/10.1038/s41537-022-00212-4

Kumar, N., Manjaly, J. a, & Sunny, M. M. (2015). The relationship between action-effect monitoring and attention capture. *Journal of Experimental*

Psychology: General, *144*(1), 18–23. https://doi.org/10.1037/xge0000032

Maeda, T., Kato, M., Muramatsu, T., Iwashita, S., Mimura, M., & Kashima, H. (2012). Aberrant sense of agency in patients with schizophrenia: Forward and backward over-attribution of temporal causality during intentional action. *Psychiatry Research*, *198*(1), 1–6. https://doi.org/10.1016/j.psychres.2011.10.021

Maniscalco, B., & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness and Cognition*, *21*(1), 422–430. https://doi.org/10.1016/j.concog.2011.09.021

Metcalfe, J., Eich, T. S., & Miele, D. B. (2013). Metacognition of agency: Proximal action and distal outcome. *Experimental Brain Research*, *229*(3), 485–496. https://doi.org/10.1007/s00221-012-3371-6

Metcalfe, J., & Greene, M. J. (2007). Metacognition of agency. *Journal of Experimental Psychology. General*, *136*(2), 184–199. https://doi.org/10.1037/0096-3445.136.2.184

Metcalfe, J., & Shimamura, A. P. (1995). *Metacognition: Knowing about knowing*. MA: MIT Press.

Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian Probability: From Neural Origins to Behavior. *Neuron*, *88*(1), 78–92. https://doi.org/10.1016/j.neuron.2015.09.039

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: A review of cue integration approaches. *Consciousness and Cognition*, *21*(1), 59–68. https://doi.org/10.1016/j.concog.2011.08.010

Nelson, T. O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. In *Psychological Bulletin* (Vol. 95, Issue 1, pp.

109–133). https://doi.org/10.1016/j.cub.2015.10.018

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect." *Psychological Science*, *2*(4), 267–270. https://doi.org/10.1109/CIS.Workshops.2007.203

Pouget, A., Drugowitsch, J., & Kepecs, A. (2016). Confidence and certainty: Distinct probabilistic quantities for different goals. *Nature Neuroscience*, *19*(3), 366–374. https://doi.org/10.1038/nn.4240

Rouault, M., Seow, T., Gillan, C. M., & Fleming, S. M. (2018). Psychiatric symptom dimensions are associated with dissociable shifts in metacognition but not task performance. *Biological Psychiatry*, *84*(6), 443–451. https://doi.org/10.1016/j.biopsych.2017.12.017

Rounis, E., Maniscalco, B., Rothwell, J. C., Passingham, R. E., & Lau, H. (2010). Theta-burst transcranial magnetic stimulation to the prefrontal cortex impairs metacognitive visual awareness. *Cognitive Neuroscience*, *1*(3), 165–175. https://doi.org/10.1080/17588921003632529

Saito, N., Takahata, K., Yamakado, H., Sawamoto, N., Saito, S., Takahashi, R., Murai, T., & Takahashi, H. (2017). Altered awareness of action in Parkinson's disease: Evaluations by explicit and implicit measures. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/s41598-017-08482-0

Salomon, R., Lim, M., Kannape, O., Llobera, J., & Blanke, O. (2013). "Self pop-out": Agency enhances self-recognition in visual search. *Experimental Brain Research*, *228*(2), 173–181. https://doi.org/10.1007/s00221-013-3549-6

Sanders, J. I., Hangya, B., & Kepecs, A. (2016). Signatures of a Statistical Computation in the Human Sense of Confidence. *Neuron*, *90*(3), 499–506. https://doi.org/10.1016/j.neuron.2016.03.025

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*(3), 241–255. https://doi.org/10.1016/j.cognition.2004.04.003

Schram, M., & Grünbaum, T. (2018). Sense of agency for movements. *Consciousness and Cognition*, *65*, 27–47. https://doi.org/10.1016/j.concog.2018.07.002

Sherman, M. T., Seth, A. K., & Barrett, A. B. (2018). Quantifying metacognitive thresholds using signal-detection theory. *BioRxiv*, 361543. https://doi.org/doi.org/10.1101/361543

Stern, Y., Ben-Yehuda, I., Koren, D., Zaidel, A., & Salomon, R. (2022). The dynamic boundaries of the Self: Serial dependence in the Sense of Agency. *Cortex*, *152*, 109–121. https://doi.org/10.1016/j.cortex.2022.03.015

Synofzik, M., Vosgerau, G., & Lindner, A. (2009). Me or not me - An optimal integration of agency cues? *Consciousness and Cognition*, *18*(4), 1065–1068. https://doi.org/10.1016/j.concog.2009.07.007

Tobler, W. R. (1965). Computation of the correspondence of geographical patterns. *Papers of the Regional Science Association*, *15*(1), 131–139. https://doi.org/10.1007/BF01947869

Tsakiris, M., Haggard, P., Franck, N., Mainy, N., & Sirigu, A. (2005). A specific role for efferent information in self-recognition. *Cognition*, *96*(3), 215–231. https://doi.org/10.1016/j.cognition.2004.08.002

Wang, S., Rajananda, S., Lau, H., & Knotts, J. D. (2021). New measures of agency from an adaptive sensorimotor task. *PLoS ONE*, *15*(12 December), 1–13. https://doi.org/10.1371/journal.pone.0244113

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious agency: Experiencing control over the movements of others. *Journal of Personality and Social Psychology*, *86*(6), 838–848. https://doi.org/10.1037/0022-3514.86.6.838

Wen, W. (2019). Does delay in feedback diminish sense of agency ? A review. *Consciousness and Cognition*, *73*(March), 102759. https://doi.org/10.1016/j.concog.2019.05.007

Wen, W., & Haggard, P. (2018). Control changes the way we look at the world. *Journal of Cognitive Neuroscience*, *30*(4), 603–619. https://doi.org/10.1162/jocn_a_01226

Wen, W., & Haggard, P. (2020). Prediction error and regularity detection underlie two dissociable mechanisms for computing the sense of agency. *Cognition*, *195*, 104074. https://doi.org/10.1016/j.cognition.2019.104074

Wen, W., & Imamizu, H. (2022). The sense of agency in perception, behaviour and human–machine interactions. *Nature Reviews Psychology*, *0123456789*. https://doi.org/10.1038/s44159-022-00030-6

Wen, W., Ishii, H., Ohata, R., Yamashita, A., Asama, H., & Imamizu, H. (2021). Perception and control: individual difference in the sense of agency is associated with learnability in sensorimotor adaptation. *Scientific Reports*, *11*(1), 1–8. https://doi.org/10.1038/s41598-021-99969-4

Wen, W., Ishikawa, T., & Sato, T. (2013). Individual differences in the encoding processes of egocentric and allocentric survey knowledge. *Cognitive Science*, *37*(1), 176–192. https://doi.org/10.1111/cogs.12005

Wen, W., Ohata, R., Tanaka, M., Yamashita, A., Asama, H., & Imamizu, H. (2019).

Two dissociable processes for detecting gaining and losing control in human brain. *The 23rd Annual Meeting of Association for the Scientific Study of Consciousness (ASSC23)*.

Wen, W., Shibata, H., Ohata, R., Yamashita, A., Asama, H., & Imamizu, H. (2020). The active sensing of control difference. *IScience*, *23*(5), 101112. https://doi.org/10.1016/j.isci.2020.101112

Wen, W., Yamashita, A., & Asama, H. (2015). The influence of goals on sense of control. *Consciousness and Cognition*, *37*, 83–90. https://doi.org/doi:10.1016/j.concog.2015.08.012