Sense of Agency for Mental Actions: Insights from a Belief-Based Action-Effect Paradigm.

Edmundo Lopez-Sola^{a,b,*}, Rubén Moreno-Bote^c, Xerxes D. Arsiwalla^{a,*}

^a*Pompeu Fabra University, Barcelona, Spain*

^b*Neuroelectrics Barcelona, Av. Tibidabo 47bis, 08035, Barcelona, Spain*

^c*Center for Brain and Cognition and Department of Information and Communications Technologies, Pompeu Fabra University, Barcelona, Spain*

Abstract

A substantial body of research in the past few decades has converged on the idea that the so-called “sense of agency”, the feeling of being in control of our own actions, arises from the integration of multiple sources of information at different levels. In this study, we investigated whether a measurable sense of agency can be detected for mental actions, without the contribution of motor components. We used a fake action-effect paradigm, where participants were led to think that a motor action or a particular thought could trigger a sound. Results showed that the high-level sense of agency, measured through explicit reports, was of comparable strength for motor and mental actions. The ‘intentional binding’ effect, a phenomenon typically associated with the experience of agency, was also observed for both motor and mental actions, with the only exception of short action-effect delays. Furthermore, a consistent relationship between explicit reports of agency and intentional binding was found. Taken together, our results provide novel insights into the specific role of intentional cues in instantiating a sense of agency, even in the absence of motor signals. These results may have important implications for future brain-computer interfaces as well as for the study of pathological disruptions of agency.

Keywords: Sense of agency, Intentional binding, Brain-computer interface, Mental action, Motor action

*Corresponding authors: edmundo.lopez@neuroelectrics.com, x.d.arswalla@gmail.com

1. Introduction

“Sense of agency” (SoA) refers to the subjective experience or feeling of being in control of our actions and their consequences. When we walk, type, or perform voluntary actions, we hold a subjective feeling of being the agents initiating those actions. How the sense of control over our actions is constructed has been the subject of a large body of research in the last twenty years, and is still a highly debated problem in cognitive science [1].

Different accounts have been put forward to explain how the sense of agency is constructed. The early years saw a dichotomy between models based on postdictive inferential processes (D. Wegner’s theory of apparent mental causation [2]) and models relying on predictive mechanisms (comparator model theories [3]). This has progressively been replaced by unifying accounts that consider both types of signals as crucial for the realization of a sense of agency [4, 5, 6]. According to these views, the SoA arises from the integration of multiple cues of distinct origins, such as motor signals, sensory feedback, background beliefs and affective cues. In fact, it has been argued that different cues could be integrated at different levels. Synofzik and colleagues [6] have proposed distinguishing between a low-level, pre-reflective feeling of agency (FoA), and a high-order, explicit judgment of agency (JoA). The FoA would arise at the sensorimotor level from the interplay of predictive signals (motor predictions, efference copy) and postdictive inferences, and corresponds to the “raw” feeling of being in control of an action. The JoA refers to explicit attributions of authorship (“I did this”) and emerges at the conceptual level, predominantly based on the low-level FoA but also influenced by contextual cues and high-order beliefs.

Despite the general consensus that sense of agency depends on the integration of multiple cues at different levels, the specific contribution of each of these cues to each level is still a matter of debate [4, 6]. In particular, the relevance of motor signals has been repeatedly emphasized, and some authors have suggested that without motor predictions the SoA would be largely reduced [7, 8, 9]. However, to date most experimental paradigms do not conclusively address the issue of whether or not the SoA could emerge in the absence of motor cues. One specific limitation of most paradigms comparing the SoA for motor and non-motor actions is that the participants’ intentionality to produce an action is usually absent in the non-motor condition [8, 9]. This is a problem, since intentionality is a crucial element in the construction of the SoA. In order to dissociate the concrete role of motor

38 predictions from other mechanisms contributing to the SoA, one would need
39 to design a paradigm where an *intentional* overt motor action is contrasted
40 with an *intentional* action that does not involve motor signals.

41 It is intuitive to think that in a situation where an intentional mental
42 event (e.g., thinking about moving an object) would be followed by the de-
43 sired outcome (the object moved), the match between intention and sensory
44 feedback would lead to the emergence of a sense of agency. Interestingly, the
45 use of Brain-Computer Interfaces (BCI) allows subjects to produce controlled
46 outcomes in the outside world, such as moving a prosthetic arm or control-
47 ling a cursor on a computer screen, through neuro-psychological processes
48 that do not necessarily involve bodily movements [10]. In the following, we
49 will refer to such neuro-psychological processes as *mental actions*. The term
50 *mental action* has been intensively discussed in the philosophy of mind and
51 action, and it is unclear which specific mental events should be regarded as
52 mental actions [11, 12]. Here, we adopt the term *mental action* for that in
53 which the cognitive capacities of an agent are recruited to produce an effect
54 of some kind, for example, “solving a chess problem in one’s head, or delib-
55 erating about whether to accept a job offer” [13]. The mental action thus
56 defined can eventually trigger outcomes in the external world through the
57 interaction with BCIs (such actions are also referred to as “BCI-mediated
58 actions”).

59 Several studies have investigated the SoA arising from the interaction with
60 BCIs [14, 15, 16, 17, 18, 19] most of them suggesting that a SoA may result
61 from BCI-mediated actions under appropriate conditions. This is coherent
62 with the aforementioned idea that intentional mental actions could lead to
63 a subjective experience of control without the presence of motor commands.
64 All these studies have assessed the SoA with explicit reports of agency, and
65 therefore, they have so far only measured high-level JoA from participants.

66 At this stage, several questions remain unanswered: First of all, it is un-
67 clear if the SoA resulting from mental actions is comparable to the strong
68 experience of agency resulting from motor actions. Secondly, if mental ac-
69 tions do lead to a SoA, as indicated by BCI studies, it is not known whether
70 this SoA emerges only at a conceptual level (JoA) from the combination of
71 thoughts, intentions and contextual cues, or if a low-level FoA also exists at
72 the non-conceptual level in the absence of motor signals.

73 In the present study, we have investigated these issues by contrasting the
74 SoA arising from motor actions with the one resulting from mental actions.
75 To that purpose, we have used an action-effect paradigm where a given mo-

76 tor or mental action was followed by a sound. In the motor action condition,
77 participants had to press a key. In the mental action condition, a “fake” BCI
78 was used: participants were led to believe that one specific thought could be
79 detected by a computer software (see “Methods”), which would subsequently
80 produce a sound. Other studies have succeeded at creating an illusion of con-
81 trol using fake BCIs [17, 16]. The paradigm is based on the idea that a causal
82 action-effect link is not necessary for the emergence of SoA. It is sufficient
83 that the action and its alleged effects are close enough in time and that
84 the subject *believes* there exists a causal relationship between both events.
85 Importantly, in order to ensure that both experimental conditions were ad-
86 equately similar in this study, no causal relationship existed in the motor
87 action condition: participants were led to think that their key press would
88 eventually produce a sound, but the sound always appeared independently
89 of the participants’ action.

90 Based on previous studies [20, 9, 21], we manipulated the time interval
91 between the action and the effect. Temporal contiguity between an action
92 and its outcome is a strong indicator of authorship, and shorter intervals
93 typically lead to a stronger SoA [22, 21]. Additionally, in order to introduce
94 a certain degree of ambiguity in the attribution of authorship, participants
95 were told that the sound could be produced either by their action or by the
96 computer.

97 We used explicit and implicit measures of the sense of agency [1]. As an
98 explicit measure, we asked participants to report whether they had produced
99 the outcome or not (authorship report). The implicit measure used was the
100 intentional binding (IB) paradigm [23]. In their seminal study, Haggard
101 and colleagues found that time perception was altered upon execution of
102 voluntary actions: the action, e.g. a key press, was perceived later in time,
103 and its effect, e.g. a sound, was perceived as occurring earlier; so that both
104 events were subjectively perceived to be closer in time. Crucially, this effect
105 was not found when the action was involuntary (a TMS pulse over the brain
106 motor area was applied to contract the muscle that pressed the key). We
107 used both types of measures to clarify at which level the SoA for mental
108 actions is constructed, explicit reports being a marker of high-level agency,
109 and the intentional binding effect being typically associated with a low-level
110 perceptual feeling of agency [24].

111 The main hypothesis of this study is that mental actions can produce a
112 sense of agency, and that this sense of agency will be comparable to that of
113 motor actions. In particular, we hypothesize that:

- 114 • The judgement of agency (JoA), measured at the conceptual level by
115 explicit reports, will be similar for mental actions to the one for motor
116 actions. We predict that for shorter action-effect intervals participants
117 will attribute the sensory effect more frequently to themselves than to
118 the computer; and that this frequency will be comparable for motor
119 and for mental actions.
- 120 • Regarding the low-level (perceptual) feeling of agency (FoA), some au-
121 thors have recently argued that predictive mechanisms based on coun-
122 terfactuals and high-order cognitive sources (intentions, beliefs, etc.),
123 different from predictive motor models, could play a key role in sev-
124 eral body-related action-outcomes, including agency attribution [25].
125 Based on this idea, and on the fact that other mechanisms such as
126 inferential postdictive processes may also contribute to the feeling of
127 agency [6], we hypothesize that a low-level FoA will arise from mental
128 actions. This is of special interest for the current research on agency,
129 since, to our knowledge, this is the first study investigating low-level
130 agency in the absence of motor cues. Our prediction is that mental
131 actions will produce intentional binding to an extent similar to that of
132 motor actions. It is worth noting that some studies have suggested that
133 the IB effect only arises when motor cues are present [9], while others
134 have argued that it is a general signature of intentionality, independent
135 of motor commands [26, 27]. Others have even claimed that it reflects
136 the mere experience of causality [28]. The present study seeks to clarify
137 whether the IB effect is specifically linked to motor actions or not.
- 138 • Finally, given that a higher sense of agency is expected for short action-
139 effect delays, our prediction is that the IB effect will be found for short,
140 but not for long intervals, and that this will be the case both for mo-
141 tor and mental actions. However, some studies using the IB paradigm
142 have not found different binding for different time intervals [24], and
143 the relationship between IB and explicit measures of agency is currently
144 unclear [29, 30, 31]. This study will offer a novel insight into the re-
145 lationship between implicit and explicit measures of agency, both with
146 and without motor cues.

147 Critically, these predictions are valid only for participants that actually
148 believe that their motor or mental actions are causally linked to the sound
149 (in reality, this causal link does not exist). Therefore, it is crucial for the

150 experiment that most participants feel that their actions can produce the
151 alleged effect. In this sense, we expect more participants to have causal beliefs
152 for the motor condition than for the mental condition, given the increased
153 difficulty of creating causal beliefs for mental actions (see “Methods”).

154 2. Methods

155 It is worth mentioning that the experiment was initially thought to be run
156 in a laboratory environment. However, due to the SARS-CoV-2 pandemic,
157 the experiment was adapted to be performed online. More details will be
158 given in the following paragraphs.

159 2.1. Participant selection

160 50 participants (23 male, 22 female, 5 unknown) were recruited through
161 the online platform Prolific (<https://www.prolific.co/>). To ensure that
162 the instructions of the experiment were correctly understood by the partici-
163 pants, only English native speakers were selected.

164 Seven participants were excluded from the study due to system incom-
165 patibilities and one participant was excluded due to poor performance in the
166 tasks.

167 2.2. Experimental design

168 2.2.1. Software tools

169 The study was designed with the open-source python-based software Psy-
170 chopy3 [32]. The experiment was translated into JavaScript and uploaded
171 to the online platform Pavlovia (<https://pavlovia.org/>). The study URL
172 from Pavlovia was provided to the participants for the execution of the ex-
173 periment in their own devices. The full python and Javascript codes for
174 the study are available at [https://gitlab.pavlovia.org/edmundolopez/](https://gitlab.pavlovia.org/edmundolopez/agency_exp)
175 `agency_exp`.

176 2.2.2. “Fake” BCI: a facial motion detection software

177 For the experiment to be successful, it was crucial to ensure that par-
178 ticipants believed that their mental actions could trigger an external event
179 through the apparent BCI. Participants were told that they would partici-
180 pate in an experiment to evaluate the users’ interaction with a facial motion
181 detection software. They were told that this software could detect micro-
182 gestures in the faces of participants, which were recorded with their device

183 webcam. They were informed that the software could recognize involuntary
184 micro-expressions in their faces when thinking about a word related to a face
185 gesture. In this case the word was ‘Smile’. During the trials, participants
186 had to think about the word ‘Smile’ at a particular moment and they were
187 told that this thought would produce a micro-gesture in their face that would
188 be detected by the software which would eventually produce a sound. Thus,
189 they would produce a sensory event with a mental action. Critically, par-
190 ticipants were instructed to avoid producing any voluntary face movement
191 during the trial.

192 *2.3. Experimental procedure*

193 Participants were first given a brief introduction to the alleged purpose
194 of the experiment: assessing user interactions with a facial motion detection
195 software. They were told that the main objective of the study was “the sense
196 of control and the sense of time while interacting with the software”.

197 The experiment consisted of three distinct blocks, one per experimental
198 condition: first, the SoA of the subjects was assessed through explicit and
199 implicit measures in a motor action task; then, the same measures were used
200 to evaluate the SoA in a mental action task; finally, in the last block we mea-
201 sured the time estimation error made by participants when not performing
202 any action (baseline condition).

203 *2.3.1. Motor action condition*

204 Participants were told that in the first block of the experiment their sense
205 of control and time over motor actions would be measured, without using
206 the facial detection software. The motor action condition was divided in
207 two parts: in the first part, participants performed three training trials and
208 twelve test trials, and after each trial the sense of agency was measured via
209 explicit report. The second part had the same number of trials, and after
210 each trial the intentional binding was measured instead.

211 The design of the trials was based on the work of Haggard and colleagues
212 on intentional binding [23]. In each trial, participants were shown a clock
213 and a white dot rotating with a period of 2.55 seconds. A red cross appeared
214 after 500 ms in a given position of the clock (which was in fact the position
215 from where the white dot started rotating), and remained on screen during
216 500 ms, after which it disappeared. Participants were instructed to press the
217 ‘space’ key when the white dot reached the position where the red cross had
218 appeared. Crucially, the time when the dot reached that position was always

the same (2.55 seconds after dot motion onset), so that the sound could be programmed to occur at varying intervals from the moment in which participants performed the action (if following the instructions properly). The red cross appeared at a different location each time, chosen randomly among the sixty ‘minute’ positions of the clock, so that participants could not habituate to the clock position where the key press would be made. Participants could nevertheless habituate to the 2.55 seconds period, and solve the task by learning the period and not looking at the dot. However, this possibility was deemed to be fairly remote given the difficulty of the task, and it was assumed that participants would need to track the dot rotation to perform the task. The trial structure is shown in Figure 1.

In every trial, the sound always appeared, regardless of when and whether the ‘space’ key was pressed. The sound appeared at one of three different intervals from the cued time: 200, 400 or 600 ms. Each interval appeared in total of four trials, and the order of appearance of each interval was randomized. Before the trials, participants were told that their key press would trigger a sound, but that in some of the trials the sound would be produced by the computer instead.

In the first part of this block, the JoA was assessed with explicit measures. Following each trial, participants were asked the question: “Who produced the tone?”. They could choose between “Me” and “The computer”. After this response, they had to respond using a Likert 1-7 scale to the following question: “How confident are you about your answer?”. The scale had in its extremes the labels “Not at all” and “Very much”. In the second part of the motor action condition, the FoA was evaluated. Participants had to indicate after each trial the clock position corresponding to the instant where they heard the sound, which allowed us to measure the intentional binding over the sound.

An important difference with other studies using similar paradigms to measure intentional binding is that in our study participants were instructed at which specific moment the mental or motor action had to be performed. The reason for this was that the experimental design required the subjects’ actions not to be causally linked with the sound. By instructing the participants to perform the action at a specific moment, the sound could be programmed to occur after that time (thus, after the action), thereby creating an illusion of causality. In brief, we assumed that if participants pressed the key at a given time and the sound appeared right after their action, they would feel that they produced the sound, even if no actual causal relation

Figure 1: Trial structure. The dot started rotating from a random position of the clock (one out of the sixty ‘minute’ positions). After 500 ms, a red cross appeared on screen, and remained for 500 ms. The red cross appears magnified for visualization purposes. The red cross appeared at the position from where the dot had started rotating, so that the dot reached the cross position always at the same time, 2.55 ms (the dot rotation period). Participants were instructed to perform an action when the dot reached the position cued by the cross: a motor action in the first experimental block, a mental action in the second one, and no action in the third. A sound appeared after a given time interval (200, 400 or 600 ms, in randomized order) after the dot reached the cued position. After the trial, participants had to either report their judgement of agency or to judge the time when they heard the sound. It is worth noting that the stimulus presented was identical in all trials of the three blocks of the experiment, with the only difference that in each block participants were instructed to perform different actions at the cued time.

257 existed between both events and even if they were “forced” to perform the
258 action at an indicated time.

259 In our experiment, only tone binding was measured, due to the fact that
260 action binding cannot be measured in the mental action condition as there
261 is no manner in which the actual time of the thought can be obtained. In
262 any case, it was expected that the execution of voluntary actions (mental
263 or motor) led to intentional binding for the sound, as has been indicated by
264 several studies [24, 33]. In fact, it has been proposed that action binding
265 and tone binding are driven by distinct mechanisms, and that they should
266 be measured separately [33, 34]

267 *2.3.2. Mental action condition*

268 Participants then completed the second block of the experiment, where
269 their SoA for mental actions was assessed. Before the real trials, several
270 instructions about the (fake) facial motion detection software were given to
271 the participants. Again, the purpose of these instructions was to increase
272 the chance that participants believed that the software could detect their
273 thoughts through their facial micro-expressions and subsequently produce a
274 sound, thereby creating an illusion of causality between their mental actions
275 and the sensory outcome.

276 Participants completed two parts of the experiment, each consisting of
277 one training trial and twelve test trials. As in the motor action condition, in
278 the first part the JoA was measured by explicit report; in the second part,
279 the FoA was measured by the intentional binding effect. Before each single
280 trial, participants were told to position their face again for the software to
281 operate properly. During each trial, participants viewed the clock and the
282 rotating dot, a cross appeared in a given position after 500 ms, and they were
283 instructed to think about the word ‘Smile’ when the dot reached the position
284 where the cross had appeared. Again, the sound appeared at varying intervals
285 from the cued time (200, 400, 600 ms, four times each in a randomized order).
286 If participants performed the task correctly, they should have perceived that
287 their thought was followed by the sound.

288 *2.3.3. Baseline condition*

289 This last block consisted of three training trials and twelve test trials. In
290 each trial, participants viewed the clock and the rotating dot, and a cross
291 appearing after one second, but they were instructed not to perform any
292 action. They were told that their webcam was not recording and that their

key press would not have any effect. The sound appeared at different intervals from the moment when the dot reached the position where the cross had appeared (200, 400 and 600 ms, four times each in a randomized order).

After each trial, participants were asked to indicate in the clock the position where they heard the sound, so that the baseline error in the estimation of the sound time could be measured. This metric is required to compute the intentional binding effect: the estimated time of the sound under voluntary actions must be compared with the estimated time of the sound in a baseline non-active condition [30, 23].

2.3.4. *Post-experimental questions and debriefing*

Once the subjects completed all three experimental conditions, they were asked two questions regarding their experience of control over the different tasks:

- “In the first block of the experiment (press key - sound), do you think that in some trials the sound was produced by your key press?”
- “In the second block of the experiment (think ‘SMILE’ - sound), do you think that in some trials the sound was produced by your thought?”

The possible answers were “Yes” and “No”. These questions are crucial for the study because they provide information about whether participants had believed or not that they could eventually produce the sound with their motor or mental actions. If a participant replied ‘Yes’ to the first or second question, this means that they thought that at least in one trial they produced the sound; therefore, the participant believed the “story” about the facial motion detection software, or about their key press producing the sound. It is critical to separate participants who had causal beliefs from those who did not because a strong sense of agency was expected only in those subjects who believed that their motor or mental actions could trigger the sound at all.

At the end of the experiment, participants were informed about the real goal of the experiment. They were told that there was no real facial motion detection software and no causal relation between their key press and the sound.

325 2.4. Data analysis

326 Analysis of the data was carried out in Python using the *statsmodels*
327 Python module for statistical models [35] and the *SciPy* Python ecosys-
328 tem [36].

329 3. Results

330 3.1. Post-experimental questionnaire

331 Out of 42 participants, 32 (76%) reported that they thought that in some
332 trials their key press produced the sound, and 15 participants (36%) reported
333 that they thought that in some trials their mental action (thinking ‘Smile’)
334 produced the sound. All participants that believed in a causal relationship
335 between their thought and the sound also believed in a causal relationship
336 between their key press and the sound. The results shown in Sections 3.2, 3.3
337 and 3.4 correspond to the data of those participants who reported believing
338 that their actions could produce the tone: 32 participants for the motor
339 action condition, and of 15 participants for the mental action condition.

340 Only data of the aforementioned participants was used because it was ex-
341 pected that if a given subject didn’t believe that their actions could produce
342 the tone, the sense of agency would not arise. Analyzing the data of subjects
343 who didn’t have causal beliefs could be misleading because some subjects
344 could have thought that there was no causal connection from the beginning,
345 while others could have realized about that during the experiment. A large
346 number of participants (50) was initially selected so that the number of par-
347 ticipants for whom causal beliefs were reported was high enough.

348 To verify the fact that the sense of agency would not be produced (or
349 in a very reduced manner) for participants without causal beliefs, we also
350 analyzed the results for the ten participants who reported believing that
351 neither their motor actions nor their mental actions could produce the tone.
352 These results are summarized in Section 3.5.

353 3.2. Behavioral performance

354 Participants’ performance in the motor action condition was evaluated
355 by measuring their key press time and comparing it with the cued time (the
356 instant when the rotating dot reached the position where the cross had ap-
357 peared). The mean error in the key press across participants was 96 ms
358 (STD = 122 ms), which means that in general participants pressed the space

bar after the cued time but not much later, and that this pattern was consistent across participants. No key press was detected in 13 trials (out of the 763 total trials in the motor action condition). Those trials were removed. Trials with outlier key press times (± 3 STD of the group mean) were also removed (11 trials, less than 2% of the total).

In the mental action task, behavioral performance could not be measured, as it was not possible to know, by its own nature, the time when participants thought about the word ‘smile’. Therefore, we had to expect that participants performed the task correctly, and we could only encourage them to do so by reminding them the instructions of the task and emphasizing the role of the webcam and the facial motion detection software.

3.3. *Explicit reports: a similar judgement of agency for motor and mental actions.*

In order to verify the hypothesis that the JoA would arise for both motor actions and mental actions, we analyzed if the number of self-attributions (responding ‘Me’ to the question “Who produced the tone?”) compared to the number of attributions to the computer was modulated by the condition (motor action versus mental action) and also by the action-effect delay.

We used the data of all trials where an explicit attribution of agency was reported. In 10 trials no report was given, and in two trials participants reported not having heard the sound. In total, data from 543 trials was available (approximately four trials per participant in each condition and interval).

We ran a binomial logistic regression where the dependent variable was authorship attribution (“Me”/“The computer”), and the main predictors were condition (motor/mental action), time interval (200/400/600 ms), and their two-way interaction. We found a significant main effect of time interval ($z = -10.026$, $p < 0.0001$), no effect of condition ($z = -1.632$, $p = 0.103$) and no interaction between condition and time interval ($z = 1.573$, $p = 0.116$). Post-hoc Chi-Square tests revealed that the number of agency self-attributions was significantly larger for short intervals than for medium intervals ($\chi^2 = 83.7$, $\text{dof} = 1$, $p < 0.0001$), and was also significantly larger for medium intervals than for long intervals ($\chi^2 = 24.2$, $\text{dof} = 1$, $p < 0.0001$). All pairwise comparisons were Bonferroni-corrected for multiple comparisons. To further assess whether there was a difference in the authorship attribution for the motor and mental conditions, we ran separate Chi-Square tests for each time interval. No significant difference was found for any of the time

Figure 2: Self-attribution decreases for increased action-effect intervals but is independent of whether the action is mental or an overt motor movement. (a) Proportion of self-attributions for each condition, groups divided by interval. The statistical significance in the comparison between groups is shown for illustrative purposes (the statistical tests have not been carried out on the proportion of self-attributions but on the number of attributions). A significant effect of time interval was found, shorter intervals leading to a higher number of self-attributions. No effect of condition and no interaction were found. * $p < 0.0001$; 'n.s.': non-significant ($p > 0.017$); Chi-Square tests, Bonferroni corrected. (b) Mean confidence rating for the attribution of agency divided by time interval, reported in a 1–7 Likert-scale. As expected, the confidence for medium intervals is significantly lower than for short and long intervals. Error bars represent 95% confidence intervals. * $p < 0.05$; 'n.s.' = non-significant ($p > 0.05$); Tukey HSD test on pairwise comparisons.

intervals (200 ms, $\chi^2 = 2.89$, $\text{dof} = 1$, $p = 0.089$; 400 ms, $\chi^2 = 0.34$, $\text{dof} = 1$, $p = 0.558$; 600 ms, $\chi^2 = 0.1$, $\text{dof} = 1$, $p = 0.752$). Figure 2a shows the proportion of agency self-attributions for the different time intervals and for the two experimental conditions.

Summarizing, participants thought that their action had produced the sound much more frequently when the interval between their action and the sound was short than when it was long. These results are in line with the prediction that the JoA is modulated by the time interval between the action and the sound. Crucially, there was no significant difference in the amount of self-attributions between motor and mental actions, indicating that the high-level sense of agency might have similar or homologous origins for both types of actions and is therefore observed in a similar manner in the presence or absence of motor commands.

409 The confidence reported by participants for their authorship attributions
 410 was also analyzed. The mean confidence rating was 5.6 (STD = 1.3) in a 1–7
 411 Likert scale ranging from ‘Not at all’ to ‘Very much’. A between-participants
 412 two-way ANOVA on confidence ratings with time interval and action type
 413 as factors revealed a significant effect of time interval ($F(2, 532) = 10.35$,
 414 $p < 0.0001$), no significant effect of action type ($F(1, 532) = 3.65$, $p = 0.056$),
 415 and no interaction ($F(2, 532) = 1.23$, $p = 0.292$). It is a matter of discus-
 416 sion whether the results of Likert scale ratings can be treated as continuous
 417 variables [37]. We have done so in the present study given that this as-
 418 sumption is widely used, specially for 1–7 scales. A Tukey HSD test showed
 419 that the confidence rating for medium intervals (400 ms) was significantly
 420 smaller than for short intervals (200 ms, $p < 0.05$) or for long intervals (600
 421 ms, $p < 0.05$). No significant difference was found between the confidence
 422 for short and long intervals ($p = 0.51$). This is an expected result because
 423 participants are more likely to think that they caused the sound for short
 424 intervals, and that the computer did so for long intervals, but they are less
 425 confident in attributing the authorship of the sound to themselves or to the
 426 computer when an intermediate interval is used. Figure 2b summarizes the
 427 confidence ratings for the different time intervals.

428 All of the analyses carried out in this section were between participants.
 429 The reason for that is that we had a different number of subjects for the
 430 action and the motor condition (see Section 3.1).

431 3.4. *Implicit measurements: intentional binding appears under motor and* 432 *mental actions.*

433 We analyzed the participants’ time estimations under three conditions:
 434 motor action, mental action and baseline (no action). In four trials, par-
 435 ticipants reported not to have heard the sound and in 20 trials no time
 436 estimation was given. We removed outlier time estimations, which corre-
 437 spond to estimations beyond ± 2.5 STD of the group mean time estimation
 438 (in average less than one trial per subject). The total number of trials was
 439 884, approximately three trials per participant per interval per condition.

440 To evaluate whether intentional binding was found in the motor action
 441 and mental action conditions, we compared the error in sound time estima-
 442 tions under each of these conditions with the error in the baseline condition.
 443 The time estimation error is calculated by subtracting the actual time of the
 444 sound from its estimated time in a given trial. For each subject, we calcu-
 445 lated the mean baseline error and the mean error under the motor or the

mental action condition. Figure 3a shows the average estimation error for all three experimental conditions. Negative errors indicate that the sound was perceived earlier than it actually occurred, and an error significantly more negative than the baseline error reflects intentional binding.

Paired samples t-tests revealed that the sound was significantly anticipated in the motor action condition compared to the baseline condition ($t(30) = -3.55, p < 0.05$), providing evidence that intentional binding was present under motor actions. Critically, we also found that there was significant binding in the mental action condition ($t(14) = -2.54, p < 0.05$), which confirms the prediction that a low-level FoA may arise from non-motor actions. Independent samples t-test showed that there was no significant difference between the estimation error in the motor action and the mental action condition ($t(44) = 0.68, p = 0.497$), suggesting that intentional binding for the sound was produced in the same manner for both types of actions. We further confirmed that the difference in the estimation error between the two conditions was not significant with a paired sample t-test using only subjects for which data was available in both conditions ($t(13) = 0.49, p = 0.636$).

The previous results correspond to time estimations for collapsed intervals. Although collapsing the data for all intervals is a common way to study intentional binding [24, 9], in this study we also wanted to assess if the time interval played a role in the intentional binding effect (as it did for the higher-level SoA). First, we conducted a repeated measures one-way ANOVA on time estimation errors in the baseline condition with time interval as factor. As expected, no significant effect of time interval was found ($F(2, 62) = 1.5, p = 0.231$), suggesting that in the baseline condition participants did similar estimation errors for the three time intervals. Therefore, we averaged the baseline error for each participant over the three intervals and subtracted it from their estimation errors in the other two conditions, thus obtaining a direct measure of binding.

A repeated measures one-way ANOVA on sound binding scores (calculated as the mean time estimation error in the action conditions minus the mean time estimation error in the baseline condition for each participant) with time interval as factor revealed a significant main effect of time interval ($F(2, 62) = 3.87, p < 0.05$). Post-hoc pairwise paired t-tests showed that the binding effect was significantly larger for short intervals than for long intervals ($t(29) = -3.31, p < 0.017$). The significance level was Bonferroni-corrected to 0.017 for multiple comparisons. This confirms the prediction that a stronger binding would appear for short intervals, in line with the re-

Figure 3: Mental and motor actions produce intentional binding over the sound, and this binding is stronger for short than for long delays. (a) Mean error in the estimation of the sound time (with respect to the actual time) for each experimental condition. Errors more negative than the baseline error indicate intentional binding. Significant binding was found for both the motor and mental action condition, indicating the presence of low-level agency. Bars represent 95% confidence intervals. $*p < 0.05$, paired sample t-tests. (b) Average binding scores for the three time delays used in the experiment, calculated as the mean estimation error of each participant in the action conditions (mental or motor) minus the mean baseline error of that participant. Shorter delays lead to higher binding, in consonance with the explicit measures of agency. Bars represent 95% confidence intervals. $*p < 0.017$, paired sample t-test, Bonferroni corrected.

sults of the explicit measures of SoA summarized in Section 3.3. The binding effect was also larger for short than for medium intervals, and for medium than for long intervals, not significantly though ($t(29) = -1.34$, $p = 0.19$; $t(29) = -1.94$, $p = 0.06$ respectively). Figure 3b displays the binding measured for the action conditions (motor and mental together) for all three time intervals.

Finally, we wanted to evaluate if, for each time interval, there was a different binding effect for the motor and for the mental actions. This allowed us to assess the relationship between implicit and explicit measures of the sense of agency: we could see if the pattern of results observed in the explicit reports (high agency for short intervals, low agency for long intervals, and no difference between motor and mental actions) is also found for intentional binding. For each time interval, we calculated every subject’s mean estimation error in each of the three experimental conditions (motor action, mental action, baseline). Figure 4 shows the average estimation error for all three experimental conditions in each time interval. For each interval, we conducted pairwise t-tests to compare the estimation errors in each action condition (motor and mental) with the baseline error.

For short intervals (200 ms), the sound was significantly anticipated in the motor action with respect to the baseline ($t(30) = -5.66$, $p < 0.0001$), but there was no significant difference between the mental action and the baseline ($t(14) = -0.71$, $p = 0.487$). Therefore, for short action-effect delays motor actions lead to intentional binding over their effects, but mental actions do not seem to produce binding. However, an independent samples t-test suggested that the difference in time estimation errors between the motor and mental condition is not significant ($t(45) = -1.21$, $p = 0.228$), while a paired sample t-test on the subject data-set that had causal beliefs in both the motor and the action condition (14 subjects) revealed a marginally significant difference in time estimation errors in the two conditions ($t(13) = -2.27$, $p = 0.04$). Given the apparently contradictory nature of this result (marginally significant or non-significant test outcomes), it must be interpreted with caution, and it will be further elaborated in the “Discussion” section.

For medium intervals (400 ms), we found that the sound was significantly anticipated with respect to the baseline for both the motor action condition ($t(30) = -3.32$, $p < 0.001$), and the mental action condition ($t(14) = -4.44$, $p < 0.0001$). This suggests that for medium intervals binding over the tone was produced by both motor and mental actions, as predicted. Furthermore,

no significant difference was found between the time estimation errors in both conditions, as shown by an independent samples t-test ($t(45) = 1.42$, $p = 0.157$) and by a paired t-test on the reduced 14 subject data-set (see above, $t(13) = 1.03$, $p = 0.323$).

For long intervals (600 ms), there was no significant difference in time estimation errors between the motor action and baseline condition ($t(28) = 0.23$, $p = 0.820$), nor between the mental action and baseline condition ($t(13) = -1.59$, $p = 0.135$). Thus, for long intervals, no binding over the tone appeared, in line with the fact that participants attributed the authorship of the sound to the computer in most of the long interval trials. A significant difference was found between the time estimation errors in the action and the motor conditions, as revealed by an independent samples t-test ($t(45) = 2.80$, $p < 0.01$) and by a paired t-test on the 14 subject data-set ($t(12) = 2.21$, $p = 0.048$). This difference suggests that participants made higher estimation errors in the mental action condition, but this is not of special relevance for the present analysis given that in none of the conditions we found significant binding.

Figure 4: Comparison of the sound time estimation error for the three experimental conditions, data divided by time interval. For short intervals, significant binding (difference with baseline error) was found only for motor actions. For medium intervals, the binding over the sound was significant for both motor and mental actions, in line with the JoA results. For long intervals, the binding was not significant for any of these two conditions, also in consonance with the explicit reports. Bars represent 95% confidence intervals. * $p < 0.01$, 'n.s.': non-significant, paired samples t-test.

539 3.5. “Non-believers”: participants who reported no causal beliefs displayed a
540 reduced sense of agency.

541 We analyzed the responses of participants who answered “No” to the
542 two questions of the post-experimental questionnaire regarding their belief
543 of causality in the motor and mental action conditions (10 subjects). We
544 assume that those participants did not believe that their actions triggered
545 the alleged effect, and that either this belief was present from the beginning
546 of the experiment or that it appeared during the experimental tasks.

547 We assessed whether the explicit reports of agency in the motor and men-
548 tal conditions were different for “non-believers” compared to participants who
549 reported causal beliefs. We ran Chi-square tests to compare the number of
550 self-attributions in each condition and each time interval for both groups.
551 We found that, for short intervals, the number of self-attributions was sig-
552 nificantly larger for believers than for non-believers, and that this was the
553 case for the motor action condition ($\chi^2 = 12.6$, dof = 1, $p < 0.001$) and
554 the mental action condition ($\chi^2 = 9.43$, dof = 1, $p < 0.01$). For medium
555 and long intervals, no significant difference was found between believers and
556 non-believers in none of the conditions (400 ms – motor action: $\chi^2 = 1.76$,
557 dof = 1, $p = 0.185$, mental action: $\chi^2 = 0.04$, dof = 1, $p = 0.841$; 600 ms –
558 motor action: $\chi^2 = 0.13$, dof = 1, $p = 0.720$, mental action: $\chi^2 = 0.008$, dof
559 = 1, $p = 0.927$). For medium and long intervals, non-believers attributed the
560 sound to the computer in most occasions, as was the case for participants
561 with causal beliefs. These results, summarized in Figure 5a, confirm that the
562 absence of causal beliefs led to a reduced (or even absent) high-level sense of
563 agency.

564 We then studied whether intentional binding could be found in the motor
565 and mental action condition for non-believers. For each subject, we com-
566 pared the mean baseline error with the mean error under the motor or the
567 mental condition. Paired samples t-tests revealed that there was no inten-
568 tional binding in the motor action condition ($t(9) = -1.29$, $p = 0.23$), nor
569 in the mental action condition ($t(9) = 0.89$, $p = 0.40$). There was also no
570 significant difference between the mean error in the motor and mental ac-
571 tion conditions, as shown by an independent samples t-test ($t(10) = -0.13$,
572 $p = 0.89$). Figure 5b illustrates these results. We conclude that the sense
573 of agency was not present at the low level for participants without causal
574 beliefs, as opposed to participants who reported believing that their actions
575 could produce the sound. In this sense, non-believers can be considered as

Figure 5: Participants without causal beliefs reported a reduced high-level sense of agency and no intentional binding. (a) Explicit reports: proportion of self-attributions for believers (*B*) and non-believers (*NB*) in each condition and interval. The statistical significance in the comparison between groups is shown for illustrative purposes (the statistical tests have not been carried out on the proportion of self-attributions but on the number of attributions). A significant difference in the number of self-attributions by believers versus non-believers was found for short intervals in both the mental and motor action condition. $*p < 0.017$; 'n.s.': non-significant ($p > 0.017$); Chi-Square tests, Bonferroni corrected. (b) Intentional binding: mean error in the estimation of the sound time (with respect to the actual time) for each experimental condition. Errors more negative than the baseline error indicate intentional binding. No significant binding was found for the motor or the mental action condition, indicating the absence of low-level agency. Bars represent 95% confidence intervals. 'n.s.': non-significant ($p > 0.05$), paired samples t-test.

576 a “control” group, proving that intentional binding is not the mere result of
577 time contiguity, but that at least the belief of causality has to be present.

578 4. Discussion

579 In the present study, we have investigated the sense of agency arising
580 from motor actions and mental actions in an action-effect paradigm, using
581 implicit and explicit measures of agency.

582 An important outcome of this study is that we have successfully man-
583 aged to reconstruct the sense of agency in subjects by using an action-effect
584 paradigm where only beliefs of participants about causality had been manip-
585 ulated prior to performing the tasks. Indeed, the stimulus was identical in
586 the three experimental conditions: a dot would rotate and after some given
587 time a sound appeared, independent of the participants’ action. However,
588 before each condition, participants were led to believe that the sound could
589 be triggered by a motor action (a key press), by a mental action (thinking
590 about a particular word), or that it would appear spontaneously (baseline
591 condition).

592 The experimental design was based on the assumption that a sense of
593 agency would be created if prior causal beliefs were sufficiently established.
594 The results of the experiment confirm this assumption: more than two thirds
595 of the total number of participants claimed to be in control of the sound in
596 the motor action block, and more than one third of the subjects thought that
597 their mental action could trigger the sound in the mental action block. The
598 fact that a sense of agency emerged for those participants purely from the
599 manipulation of prior causal beliefs is a critical success of this experiment.
600 This was further confirmed by analyzing the results for subjects who did not
601 have causal beliefs: the sense of agency was reduced or even absent for these
602 participants.

603 4.1. High- and low-level agency

604 We assessed the high-level JoA by measuring the number of times that
605 a participant would attribute the sound to his/her own action or to the
606 computer, and comparing it for motor and mental actions. We validated the
607 finding that shorter intervals lead to higher agency under motor actions, as
608 confirmed by the higher number of self-attributions for short action-effect
609 delays, and the progressive reduction of self-attributions for medium and
610 large delays. Interestingly, this same pattern was found under mental actions,

and no significant difference was found in the number of self-attributions for mental and motor actions, confirming our main hypothesis. This result offers a novel insight to the possibility that a strong SoA arises under non-motor actions, at least at a high level (judgement of agency). Although it has already been suggested by previous studies involving BCIs that a SoA might emerge from mental actions [14, 19], our results show not only that mental actions produce a JoA, but that they do so to a similar extent as motor actions (that is, the number of self-attributions ascribed to mental actions was not significantly different from that of motor actions, in any of the above time intervals); and that the SoA for mental actions disappears progressively under longer action-effect time delays, as occurs for motor actions. Within the framework of the cue integration theory of the SoA [4, 6], these results indicate that in the absence of motor signals, other cues such as prior beliefs or non-motor predictions may contribute to the construction of judgements of agency, compensating for the lack of motor cues. This contradicts other views that propose that motor commands are necessary for a robust sense of agency to emerge, as suggested by accounts based on the comparator model [38].

We also measured the intentional binding over the sound produced by motor and mental actions, as a marker of low-level agency. We replicated the finding that voluntary motor actions produce binding over the effect of those actions: the sound was perceived as occurring earlier than it actually occurred. Critically, we also found that mental actions produced binding over the sound, and that there was no significant difference in the magnitude of the intentional binding effect for motor and for mental actions, once again confirming our hypothesis. This is a major finding of the present study, from which two separate conclusions can be extracted. First, that the intentional binding effect is not to be exclusively linked to motor actions as some authors suggest [39, 34], but is rather a signature of low-level agency for intentional actions in general, as proposed by others [26]. It must be noted that our results are also consistent with the hypothesis put forward by Buehner [28] that the IB effect could reflect pure causality and not intentionality. In any case, both accounts agree in the fact that motor cues are not exclusively required for IB to occur. Second, we conclude that mental actions give rise to a sense of agency not only at a higher conceptual level, but also at a lower perceptual level. This, we claim, is a novel finding in the research on agency. To the best of our knowledge, the current study is the first one to investigate the FoA under non-motor actions.

649 We contemplate that the construction of the FoA from mental actions
650 presumably relies on general predictive mechanisms, such as counterfactuals,
651 combined with inferential postdictive mechanisms not related to motor
652 cues [6]. In fact, it could be that the FoA for both motor and mental ac-
653 tions is in part based on cognitive predictions independent of motor-based
654 forward models, as suggested by Dogge and colleagues [25]. Based on the
655 pre-activation account of intentional binding proposed by Waszak and col-
656 leagues [39], we hypothesize that these general cognitive predictions could
657 pre-activate the perceptual representation of the sensory effect (in this case,
658 the sound), leading to intentional binding (the original proposal by Waszak
659 and colleagues suggested that these predictions were motor predictions).

660 We also investigated whether a consistent relationship between explicit
661 and implicit measures of agency could be found: if this was the case, the
662 intentional binding effect should be substantially larger for shorter intervals,
663 given that in most short-delay trials participants attributed the authorship
664 of the sound to themselves. Indeed, in the motor action condition we found
665 significant binding for short and medium intervals but no binding for large
666 intervals. This confirms that there is a coherent relationship between the
667 explicit reports of the participants and the implicit measurement of the sense
668 of agency through the intentional binding paradigm. This result is of special
669 relevance because other studies investigating the relationship between explicit
670 and implicit measures of agency have not found a fully consistent relationship
671 between both measures, concluding that they might be driven by dissociable
672 mechanisms [29, 31]. Here instead, we have found that the intentional binding
673 over the sound is a good predictor of the explicit agency reported by the
674 participants. This provides further evidence that the low-level feeling of
675 agency is the major basis for the construction of a higher level judgement of
676 agency, as hypothesized by Synofzik and colleagues [6].

677 This consistent relationship between implicit and explicit measures also
678 existed for mental actions, in particular for medium and long intervals (bind-
679 ing was found for medium delays but not for long delays). However, the bind-
680 ing over the sound was not significant for short intervals, which was not ex-
681 pected because participants had reported a high amount of self-attributions
682 for short delays in the mental action condition. This result can be inter-
683 preted in two alternative ways: either no low-level feeling of agency emerged
684 for non-motor actions under short delays, or a feeling of agency emerged but
685 was not captured by the intentional binding effect. We support the latter
686 option, for the reasons that follow. First, the high-level judgements of agency

for 200 ms intervals were higher than for medium intervals. It is widely accepted that the judgements of agency, even if influenced by other factors, are largely based on the low-level feeling of agency. If no FoA arose for 200 ms, but did so for 400 ms, one would expect higher JoA for medium than for short intervals; but this was not the case. Second, the absence of intentional binding for mental actions under short intervals can be explained by the fact that the outcomes of mental actions might be predicted slower than the outcomes of motor actions. Indeed, in our life-long experience interacting with the external environment generally requires bodily movements (BCI are an exception to this). Thus, it is coherent that the predictive mechanisms for the sensory outcomes of motor movements operate at fast timescales (fast arm movements may last less than 200 ms) [40], while the predictions of sensory events triggered by mental actions, which are less common, may be less efficient. In the framework of the pre-activation account of IB [39], this could mean that the predictive mechanisms for mental actions were not fast enough to pre-activate the perceptual representation before the stimulus arrived, and therefore no intentional binding was found. In the absence of an efficient sensory prediction, the FoA might have been constructed through other mechanisms, such as postdictive inferential processes based on the external sensory input and on cognitive priors, as proposed by cue integration theories. Nevertheless, the results of the current study do not allow us to draw a definite conclusion with respect to this matter, and it shall be further investigated whether a feeling of agency may arise for mental actions and short action-effect intervals. For that purpose, other implicit measures such as sensory attenuation paradigms may be useful [41, 42].

In summary, we have found that the sense of agency emerging at the high conceptual level was as strong for non-motor mental actions as for motor actions, and followed a similar pattern for different time intervals. At the low perceptual level, mental actions produced intentional binding in a similar manner as for motor actions, with the exception of short action-effect delays, a matter which will need further investigation. Furthermore, we have found a consistent relationship between the explicit reports and the intentional binding effect, suggesting that this phenomenon might be driven by general predictive mechanisms not limited to motor predictions. Finally, we have found that, in the absence of causal beliefs, the JoA is substantially reduced and the intentional binding effect is not found, suggesting an absence of FoA. This emphasizes the key role of action-effect causality beliefs, which are necessary for the resulting sense of agency.

725 4.2. *Limitations and additional considerations*

726 This study presents several challenges and limitations that might be ad-
 727 dressed in future work. First, one could argue that the alleged non-motor
 728 actions involve motor signals to a certain extent, because participants are
 729 told that the software detects facial micro-expressions. Even if participants
 730 are told not to produce any voluntary gesture, it is possible that some motor
 731 commands are involved in the execution of the action, for example by forc-
 732 ing a ‘static’ facial expression to avoid producing voluntary gestures. Our
 733 prediction is that the results would have been very similar had this exper-
 734 iment been carried out with a BCI that did not involve any kind of motor
 735 signal; we surmise that any possible contribution of motor cues to the mental
 736 action condition is negligible. Another limitation is that many participants
 737 did not believe that they were in control in the mental action condition. We
 738 think that using a more credible paradigm like an EEG-based BCI would
 739 greatly increase the number of participants for whom we could manipulate
 740 prior causal beliefs.

741 A potential limitation of this study is that in usual SoA experiments,
 742 participants get to choose when to perform an action, which action to perform
 743 or whether to act or not. In our everyday life, the conjunction of these three
 744 components (what to do, when to do it and whether to do it or not) gives
 745 rise to a strong feeling of intentionality in our actions. The relevance and
 746 neuroscientific evidence of the ‘what’, ‘when’ and ‘whether’ components of
 747 intentional action have been reviewed elsewhere [43]. In the present study,
 748 participants were told which action to perform and when to perform it (and
 749 following the instructions did not leave a choice whether or not to perform the
 750 action). However, we argue that the decision to follow the instructions was
 751 in itself an intention to act that participants executed, and the presence of
 752 this intentionality led to a partial version of the ‘what’, ‘when’ and ‘whether’
 753 components. This justifies the fact that participants reported a SoA under
 754 our experimental paradigm. We hypothesize that under a real BCI paradigm,
 755 where participants could choose when to act or which action to perform, the
 756 sense of agency would be stronger than the one measured in this study. This
 757 line of research remains an interesting possibility for future work.

758 It has been recently argued that the sense of agency measured in cog-
 759 nitive neuroscience is ambiguous between different constructs: the sense of
 760 agency as a cognitive ability or as a phenomenal character, and bodily ver-
 761 sus external constructs of agency. These distinctions yield a total of four
 762 distinct constructs [44]. The authors suggest that the different experimental

763 paradigms used to study the SoA are related to different constructs. We pro-
764 pose that the present study focuses on the external construct of agency, given
765 that in the tasks participants were requested to make judgements about an
766 external event. Moreover, our experiment targets both the cognitive ability
767 and the phenomenal character of SoA, since we have used the self-other dis-
768 tinction for the authorship of external outcomes (associated with the SoA
769 as a cognitive ability), and the intentional binding effect, which is related
770 to the phenomenal character of the SoA. In fact, given that we have found
771 a coherent relationship between these measures (self-other distinction and
772 intentional binding), we argue that both constructs of the SoA arise from
773 related mechanisms, and perhaps could be studied interchangeably, at least
774 when referring to the external construct of the SoA.

775 *4.3. Implications for BCI research*

776 The findings of this study have important implications for the field of
777 brain-computer interfaces. Previous studies have already presented evidence
778 that subjects performing BCI-mediated tasks report being in control of those
779 actions [14, 19]. The results presented in our study suggest that this sense
780 of agency does not exclusively arise at a pure conceptual level, but also at
781 a low perceptual level, similarly to the SoA for bodily movements. This
782 implies that interaction with BCIs may help reconstruct a strong sense of
783 control, making them promising technologies for the future in a wide range
784 of applications. In particular, the robust feeling of agency evidenced in our
785 study argues in favor of the use of BCIs for the control of external devices,
786 such as prosthetic limbs [45], and for other applications requiring a strong
787 sense of control, such as communication through brain interfaces [46].

788 Importantly, the high SoA measured for non-motor mental actions indi-
789 cates that external devices could be controlled in a robust manner not only
790 by BCIs based on motor commands (typically through motor imagery), but
791 also by non-motor-based BCIs, such as steady state visually evoked poten-
792 tial (SSVEP) based BCIs. Nevertheless, it is possible that sensorimotor BCIs
793 provide more robust solutions for applications requiring precise and fast con-
794 trol, as suggested by the absence of intentional binding for mental actions
795 and short delays in this experiment, and as also seen in other studies [19].

796 Another interesting implication of our study concerns how legal attribu-
797 tions of moral responsibility may have to be rethought in the presence of
798 future BCI technologies. Several authors have already raised issues concern-
799 ing moral responsibility over BCI-mediated actions: users might be uncertain

about their agency in a BCI context, making difficult the attribution of moral (and therefore, legal) responsibility [47, 48]. Moreover, in most legal systems this attribution of responsibility is based on the fact that only “voluntary acts” are punishable, an “act or action” being defined as a “bodily movement” (USA Model Penal Code, Section 2.01). Reflexes and unconscious movements are qualified as involuntary and, thus, are not punishable. However, BCI-mediated actions would not be proper actions under the definition of “act”, even if those are voluntary. The current findings support the idea that BCI-mediated actions should presumably be considered voluntary actions, given that subjects have a sense of agency for mental actions to a comparable extent and execution as they do for motor actions. Thus, we argue that legal systems may need to be adapted to include wider definitions of agency, notably with respect to mentally-triggered external events. The manner in which the notion of agency might be extended to mental actions in the framework of current legal systems is still a matter of debate and will be of critical relevance for the future of BCI technologies in society [47, 49, 48, 18].

4.4. *Clinical applications*

Our experiment suggests that the presence of causal beliefs may induce an illusion of control over external events which are in fact unrelated to the subjects’ actions. Such illusions of control could be used alongside the treatment of psychological disorders typically associated with a loss of control, as is often the case for patients suffering from depression or anxiety [50]. For instance, a recent study has shown that simply activating memories of depression in healthy subjects caused an increase in their estimates of intentional binding compared to activation of baseline memories, thus suggesting a partial loss of sense of agency during depressive episodes [51]. The results of our study suggest that implementing positive causal beliefs during therapy may play an important role in increasing the sense of agency in patients with depression.

Additionally, BCI-based experiments such as the one used in this study might greatly help understand pathological disruptions of the sense of agency, in particular the phenomenon of thought insertion in schizophrenia [52]. Voss and colleagues have explored the mechanisms behind the abnormal experiences of agency exhibited by schizophrenic patients through an IB paradigm [53]. Analogous to that, here we propose that investigating the sense of control over mentally-triggered external events in patients with thought insertion symptomatology may help elucidate mechanisms disrupted

under this pathology. It has been suggested that the authorship of thoughts is not a feeling arising at a low level, but rather at a high conceptual level [52]. Thus, one would expect to find a similar intentional binding (low level) for mentally-triggered events in patients suffering from thought insertion compared to healthy subjects; but schizophrenic patients would probably display abnormal attributions of authorship (high level) for the outcomes of these mental actions. This line of research remains to be explored in future work.

Acknowledgements

RM-B is supported by BFU2017-85936-P and FLAGERA-PCIN-2015-162-C02-02 from MINECO (Spain), the Howard Hughes Medical Institute (HHMI; ref 55008742), The Bial Foundation (grant number 117/18) and ICREA Academia (2016).

References

- [1] J. W. Moore, What Is the Sense of Agency and Why Does it Matter?, *Frontiers in Psychology* 7 (2016). Publisher: Frontiers.
- [2] D. M. Wegner, T. Wheatley, Apparent mental causation. Sources of the experience of will, *The American Psychologist* 54 (1999) 480–492.
- [3] C. D. Frith, S.-J. Blakemore, D. M. Wolpert, Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action, *Brain Research Reviews* 31 (2000) 357–363.
- [4] J. W. Moore, P. C. Fletcher, Sense of agency in health and disease: A review of cue integration approaches, *Consciousness and Cognition* 21 (2012) 59–68.
- [5] H. C. Lau, R. D. Rogers, R. E. Passingham, Manipulating the experienced onset of intention after action execution, *Journal of Cognitive Neuroscience* 19 (2007) 81–90.
- [6] M. Synofzik, G. Vosgerau, M. Voss, The experience of agency: An interplay between prediction and postdiction, *Frontiers in Psychology* 4 (2013). Place: Switzerland Publisher: Frontiers Media S.A.

- 866 [7] C. Frith, The self in action: Lessons from delusions of control, *Con-*
867 *sciousness and Cognition* 14 (2005) 752–770.
- 868 [8] A. Sato, Both motor prediction and conceptual congruency between pre-
869 *view and action-effect contribute to explicit judgment of agency, Cog-*
870 *nition* 110 (2009) 74–83.
- 871 [9] J. W. Moore, D. M. Wegner, P. Haggard, Modulating the sense of agency
872 *with external cues, Consciousness and Cognition* 18 (2009) 1056–1064.
- 873 [10] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H.
874 *Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, T. M.*
875 *Vaughan, Brain-computer interface technology: a review of the first*
876 *international meeting, IEEE transactions on rehabilitation engineering:*
877 *a publication of the IEEE Engineering in Medicine and Biology Society*
878 *8 (2000) 164–173.*
- 879 [11] L. O’Brien, M. Soteriou, *Mental actions*, 2010. Pages: 296.
- 880 [12] T. Metzinger, The Problem of Mental Action, in: T. Metzinger,
881 *W. Wiese (Eds.), Philosophy and Predictive Processing*, 2017.
- 882 [13] A. R. Mele, Agency and Mental Action, *Philosophical Perspectives* 11
883 (1997) 231–249. Publisher: [Ridgeview Publishing Company, Wiley].
- 884 [14] N. Evans, S. Gale, A. Schurger, O. Blanke, Visual Feedback Dominates
885 *the Sense of Agency for Brain-Machine Actions, PLOS ONE* 10 (2015)
886 *e0130019. Publisher: Public Library of Science.*
- 887 [15] M. Slater, D. Pérez Marcos, H. Ehrsson, M. V. Sanchez-Vives, Inducing
888 *illusory ownership of a virtual body, Frontiers in Neuroscience* 3 (2009).
889 *Publisher: Frontiers.*
- 890 [16] R. Vlek, J.-P. van Acken, E. Beursken, L. Roijendijk, P. Haselager, BCI
891 *and a User’s Judgment of Agency, in: G. Grübler, E. Hildt (Eds.), Brain-*
892 *Computer-Interfaces in their ethical, social and cultural contexts, The*
893 *International Library of Ethics, Law and Technology, Springer Nether-*
894 *lands, Dordrecht, 2014, pp. 193–202.*
- 895 [17] M. T. Lynn, C. C. Berger, T. A. Riddle, E. Morsella, *Mind control?*
896 *Creating illusory intentions through a phony brain–computer interface,*
897 *Consciousness and Cognition* 19 (2010) 1007–1012.

- 898 [18] H. Limerick, D. Coyle, J. W. Moore, The experience of agency in human-
899 computer interactions: a review, *Frontiers in Human Neuroscience* 8
900 (2014). Publisher: Frontiers.
- 901 [19] B. Nierula, B. Spanlang, M. Martini, M. Borrell, V. V. Nikulin, M. V.
902 Sanchez-Vives, Agency and responsibility over virtual movements controlled
903 through different paradigms of brain-computer interface, *The Journal of Physiology* (2019).
904
- 905 [20] K. Engbert, A. Wohlschläger, R. Thomas, P. Haggard, Agency, subjective
906 time, and other minds, *Journal of Experimental Psychology: Human Perception and Performance* 33 (2007) 1261–1268. Place: US
907 Publisher: American Psychological Association.
908
- 909 [21] A. Sato, A. Yasuda, Illusion of sense of self-agency: Discrepancy between
910 the predicted and actual sensory consequences of actions modulates the
911 sense of self-agency, but not the sense of self-ownership, *Cognition* 94
912 (2005) 241–255.
- 913 [22] V. Chambon, J. W. Moore, P. Haggard, TMS stimulation over the inferior
914 parietal cortex disrupts prospective sense of agency, *Brain Structure and Function* 220 (2015) 3627–3639.
915
- 916 [23] P. Haggard, S. Clark, J. Kalogeras, Voluntary action and conscious
917 awareness, *Nature Neuroscience* 5 (2002) 382–385. Number: 4 Publisher:
918 Nature Publishing Group.
- 919 [24] A. Desantis, C. Roussel, F. Waszak, On the influence of causal beliefs on
920 the feeling of agency, *Consciousness and Cognition* 20 (2011) 1211–1220.
- 921 [25] M. Dogge, R. Custers, H. Aarts, Moving Forward: On the Limits of
922 Motor-Based Forward Models, *Trends in Cognitive Sciences* 23 (2019).
- 923 [26] A. Desantis, G. Hughes, F. Waszak, Intentional Binding Is Driven by
924 the Mere Presence of an Action and Not by Motor Prediction, *PLOS ONE* 7 (2012) e29557. Publisher: Public Library of Science.
925
- 926 [27] M. Synofzik, G. Vosgerau, A. Lindner, Me or not me – An optimal
927 integration of agency cues?, *Consciousness and Cognition* 18 (2009)
928 1065–1068.

- 929 [28] M. J. Buehner, Understanding the Past, Predicting the Future: Cau-
930 sation, Not Intentional Action, Is the Root of Temporal Binding, Psy-
931 chological Science 23 (2012) 1490–1497. Publisher: SAGE Publications
932 Inc.
- 933 [29] J. P. Ebert, D. M. Wegner, Time warp: Authorship shapes the perceived
934 timing of actions and events, Consciousness and Cognition 19 (2010)
935 481–489.
- 936 [30] J. W. Moore, S. S. Obhi, Intentional binding and the sense of agency:
937 a review, Consciousness and Cognition 21 (2012) 546–561.
- 938 [31] K. A. Schwarz, L. Weller, A. L. Klaffehn, R. Pfister, The effects of
939 action choice on temporal binding, agency ratings, and their correlation,
940 Consciousness and Cognition 75 (2019) 102807.
- 941 [32] J. Peirce, J. R. Gray, S. Simpson, M. MacAskill, R. Höchenberger,
942 H. Sogo, E. Kastman, J. K. Lindeløv, PsychoPy2: Experiments in
943 behavior made easy, Behavior Research Methods 51 (2019) 195–203.
- 944 [33] N. Wolpe, P. Haggard, H. R. Siebner, J. B. Rowe, Cue integration
945 and the perception of action in intentional binding, Experimental Brain
946 Research 229 (2013) 467–474.
- 947 [34] N. Wolpe, J. B. Rowe, Beyond the “urge to move”: objective mea-
948 sures for the study of agency in the post-Libet era, Frontiers in Human
949 Neuroscience 8 (2014). Publisher: Frontiers.
- 950 [35] S. Seabold, J. Perktold, Statsmodels: Econometric and Statistical Mod-
951 eling with Python, Proceedings of the 9th Python in Science Conference
952 2010 (2010).
- 953 [36] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy,
954 D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J.
955 van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J.
956 Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, I. Polat, Y. Feng,
957 E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman,
958 I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H.
959 Ribeiro, F. Pedregosa, P. van Mulbregt, SciPy 1.0: fundamental algo-
960 rithms for scientific computing in Python, Nature Methods 17 (2020)
961 261–272. Number: 3 Publisher: Nature Publishing Group.

- [37] J. Carifio, R. J. Perla, Ten Common Misunderstandings, Misconceptions, Persistent Myths and Urban Legends about Likert Scales and Likert Response Formats and their Antidotes, *Journal of Social Sciences* 3 (2007) 106–116.
- [38] C. D. Frith, Action, agency and responsibility, *Neuropsychologia* 55 (2014) 137–142.
- [39] F. Waszak, P. Cardoso-Leite, G. Hughes, Action effect anticipation: Neurophysiological basis and functional consequences, *Neuroscience & Biobehavioral Reviews* 36 (2012) 943–959.
- [40] R. C. Miall, D. M. Wolpert, Forward Models for Physiological Motor Control, *Neural Networks* 9 (1996) 1265–1279.
- [41] S. J. Blakemore, D. Wolpert, C. Frith, Why can’t you tickle yourself?, *Neuroreport* 11 (2000) R11–16.
- [42] G. Hughes, A. Desantis, F. Waszak, Mechanisms of intentional binding and sensory attenuation: The role of temporal prediction, temporal control, identity prediction, and motor prediction, *Psychological Bulletin* 139 (2013) 133–151. Place: US Publisher: American Psychological Association.
- [43] M. Brass, P. Haggard, The What, When, Whether Model of Intentional Action:, *The Neuroscientist* (2008). Publisher: SAGE PublicationsSage CA: Los Angeles, CA.
- [44] T. Grünbaum, M. S. Christensen, Measures of agency, *Neuroscience of Consciousness* 2020 (2020). Publisher: Oxford Academic.
- [45] J. Clausen, E. Fetz, J. Donoghue, J. Ushiba, U. Spörhase, J. Chandler, N. Birbaumer, S. R. Soekadar, Help, hope, and hype: Ethical dimensions of neuroprosthetics, *Science* 356 (2017) 1338–1339. Publisher: American Association for the Advancement of Science Section: Policy Forum.
- [46] C. Grau, R. Ginhoux, A. Riera, T. L. Nguyen, H. Chauvat, M. Berg, J. L. Amengual, A. Pascual-Leone, G. Ruffini, Conscious Brain-to-Brain Communication in Humans Using Non-Invasive Technologies, *PLOS ONE* 9 (2014) e105225. Publisher: Public Library of Science.

- 993 [47] S. Steinert, C. Bubnitz, R. Jox, O. Friedrich, Doing Things with
994 Thoughts: Brain-Computer Interfaces and Disembodied Agency, *Phi-*
995 *losophy & Technology* 32 (2019) 457–482.
- 996 [48] P. Haselager, Did I Do That? Brain–Computer Interfacing and the
997 Sense of Agency, *Minds and Machines* 23 (2013) 405–418.
- 998 [49] J. M. Vukov, K. Rempala, BCI-Mediated Action, Blame, and Responsi-
999 bility, *AJOB Neuroscience* 11 (2020) 65–67. Publisher: Taylor & Francis
1000 _eprint: <https://doi.org/10.1080/21507740.2019.1704929>.
- 1001 [50] C. P. Keeton, M. Perry-Jenkins, A. G. Sayer, Sense of Control Predicts
1002 Depressive and Anxious Symptoms Across the Transition to Parenthood,
1003 *Journal of family psychology : JFP : journal of the Division of Family*
1004 *Psychology of the American Psychological Association (Division 43)* 22
1005 (2008) 212–221.
- 1006 [51] S. S. Obhi, K. M. Swiderski, R. Farquhar, Activating memories of de-
1007 pression alters the experience of voluntary action, *Experimental Brain*
1008 *Research* 229 (2013) 497–506.
- 1009 [52] G. Vosgerau, M. Voss, Authorship and Control over
1010 Thoughts, *Mind & Language* 29 (2014) 534–565. _eprint:
1011 <https://onlinelibrary.wiley.com/doi/pdf/10.1111/mila.12065>.
- 1012 [53] M. Voss, J. Moore, M. Hauser, J. Gallinat, A. Heinz, P. Haggard, Al-
1013 tered awareness of action in schizophrenia: a specific deficit in predicting
1014 action consequences, *Brain* 133 (2010) 3104–3112. Publisher: Oxford
1015 Academic.