Agency Attribution and Temporal Binding. Towards a New Model of Time Perception.

Beatrice Kaloczy[1[0000-0002-1200-3124]]

[1] University College Dublin, Belfield, Dublin 4, Ireland
beatrice.kaloczy@ucdconnect.ie

Abstract. Sense of agency is a concept which represents our ability to attribute actions to ourselves or other people. Studies found that when we perform self-generated actions and when we observe other people doing the same, a phenomenon known as temporal binding occurs. It consists of a perception of "temporal attraction" between actions and outcomes [5,8]. Some studies proved that agency and temporal binding rely on multiple top-down and bottom-up mechanisms, such as inferential and sensorimotor processes [18]. However, not many researches considered the impact of experimental factors. For this reason, the present study will focus on this aspect, particularly on the role of the interface and the measure of agency, in order to shed more light on the mechanisms underlying agency attribution and temporal binding. From one side, we found no significant difference in temporal binding in screen and virtual reality conditions, although virtual reality seemed to increase the accuracy of time perception. From the other side, we found that the range of intervals significantly influenced people's estimations, and errors linearly diverged from one specific interval which was estimated with the most accuracy. Our hypothesis is that people use contextual cues and their knowledge of the world to find an optimal interval between actions and outcomes in different circumstances. This not only questions what we know about agency and temporal binding, but it can be also the foundation of a new model of time perception, where long-term memory and context play a fundamental role.

Keywords: Agency Attribution, Virtual Reality, Time Perception.

1. Introduction

An important part of our existence is related to the perception of control we have over our actions and the world around us [5,7]: this concept is called sense of agency [5,7]. In more detail, we normally experience sense of agency when we perform an action and we clearly identify that action as initiated by us and no one else [16]. Sense of agency can be divided into two components: "feeling of agency" and "judgment of agency" [16]. The former consists of a pre-reflective and implicit feeling of control, whereas the latter is "conscious" and can be expressed by words [12,16]. Thus, from one side, we immediately and implicitly register whether an action is initiated by us or not [16], from the other side, we can consciously evaluate the circumstances and attribute actions to ourselves [16]. Thus, agency seems to be an extremely multifaceted concept, which arises from various mechanisms. From one side, it is connected to bottom-up processes based on multi-sensory feedbacks and proprioception [17]: we feel in control of our actions when we detect congruency between actual and predicted outcomes [17]. From the other side, agency is influenced by inferential top-down processes based on contextual information, individual experience and personal knowledge [17]: these processes allow to elaborate "non-conceptual" information coming from the sensorimotor system, and to explicitly identify ourselves as the cause of our own actions [17].

A component of sense of agency is called agency attribution, which is our ability to attribute actions to other people. It turned out that agency for self-generated actions and agency attribution seem to rely on similar neural and cognitive processes [12]. For example, Poonian and colleagues (2015) [12] found that when we perform self-generated actions and when we observe other people, there is no significant difference in the suppression of the ERP component N1, traditionally associated with self-generated actions which produce auditory or visual outcomes [12]. This seems to be consistent also with a behavioral phenomenon known as intentional binding [12], often used to measure feeling of agency. Traditionally, intentional binding is considered as a perceived "temporal attraction" between "voluntary" actions and their outcomes [5,8], but it turned out to be present also when we observe other people or when we participate in joint actions with them [10,12]. Methods which have often been used to measure binding effects are for example the Libet clock (Figure 1a) and the interval estimation procedure (Figure 1b), where people have to respectively report when actions and outcomes occur or estimate the interval between them [3,8].

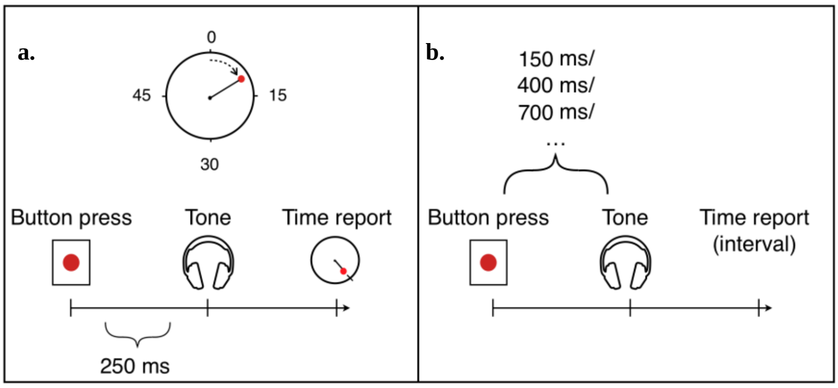

Figure 1. Graphical representation of the Libet clock and interval estimation procedure. In both cases, participants are traditionally asked to press a button which gives rise to a tone. **a.** The Libet clock is presented on a computer screen and its hand rotates and completes a cycle in 2560 ms [3]. Participants need to report where the hand of the clock was located either when they pressed the button or when they heard the tone [3]. **b.** In the interval estimation procedure,

participants need to estimate the interval of time elapsed between the action and its auditory outcome which varies across trials [3]. In both methods, people tend to perceive their voluntary action and its outcome as temporally closer [3]. This phenomenon is called temporal binding.

However, researchers interpret the mechanisms underlying temporal binding very differently. Particularly, some studies about agency attribution conducted in the field of Human Computer Interaction leaded to two different approaches. From one side, researchers believe that temporal binding and agency are related to intentionality. One of the most famous studies in this area was conducted by Obhi & Hall (2011) [11], who found that intentional binding was present only in human-human interactions and not in joint actions with computers. To obtain these results, it was sufficient to manipulate people's belief and tell them that they were interacting either with a person or a computer [11]. According to the authors, simply knowing that we are interacting with other humans gives rise to a "pre-reflective we agentic identity" [11]: in fact, we can attribute agency to other people, predict and represent their actions and their intentions as they were our own, because we biologically share similar ways of thinking and behaving, [11,14,15]. Probably, mirror neurons are involved, as they are active not only when we perform certain tasks but also when we observe other people [13].

From the other side, other researchers found that temporal binding is present also in human-machine interactions, and believe that this phenomenon reflects causal effects rather than intentionality [2]. For example, Buehner (2012) [2] found temporal binding not only when participants observed other people performing actions, but also when they observed a mere "mechanical causation". Thus, according to the author, temporal binding is mainly connected to causation between actions and their effects [2]. In this case, temporal binding can be explained by the "theory of Bayesian ambiguity reduction" [2]: when two events are causally linked, they are also more likely to be temporally close to each other [2]. For this reason we may experience a temporal attraction between causal events when we are uncertain about the interval which elapses between them [2].

A model which seems to explain these inconsistent findings is the Optimal Cue Integration model proposed by Synofzik and colleagues (2013) [18]. According to these authors, cues such as sensorimotor causal effects, predictions, feedbacks, beliefs, knowledge about context, previous experience, or even "affective" components, interact between each other, and are weighted and used differently depending on the "availability" and "reliability" of the information [18]. Thus, agency and temporal binding cannot be explained by a single mechanism such as intentionality or causality, but rather by a complex interconnection of bottom-up and top-down processes, where different cues are relevant in different situations. Although some researches proved that multiple processes and cues influence sense of agency [4], there are some factors which have not been investigated yet. Particularly, we think that specific experimental conditions can dramatically influence sense of agency and binding effects.

The aim of this project is to explore some of these aspects and question the traditional means and procedures used in the study of agency. The present experiment is based on more ecological and perceptually complex stimuli, compared to traditional agency attribution tasks. In fact, participants will be presented with an archery scenario and they will need to estimate the interval between the time when the archer shoots the arrow and when the arrows reaches the target. There will be two agent conditions: in one case, participants will be told that the archer is controlled by a computer, in the other case, they will be told that it is controlled by another person. In this scenario, two different interfaces will be directly compared for the first time in this field: a traditional computer screen and virtual reality. In addition to that, it will be investigated whether time perception and binding effects are influenced by changes in the measure of agency, in this case the interval estimation procedure. Particularly, it will be used a wider range of intervals compared to previous studies. In the next sections, the experiment will be outlined and all these aspects will be addressed in order to shed more light on the mechanisms underlying agency attribution and temporal binding.

2. Methods

2.1. Participants

28 participants (16 females and 12 males) took part in the experiment and were selected by following these selection criteria: normal or corrected to normal vision and no self-reported auditory or motor impairments. Participants were aged between 21 and 36 years old, with a general mean of 25 years old. Most of the participants were students or employees at University College Dublin and were invited to reach the Visualization Cave in the Insight Centre for Data Analytics located in the same university. The project received ethical approval from the Human Research Ethics Committee - Sciences (HREC-LS) of University College Dublin.

2.2. Experimental design

The experiment had a 2x2 matrix design, where interface and type of agent were the two independent variables with two levels each. The interfaces taken in consideration were virtual reality (HTC Vive headset) and a traditional screen (13-inch MacBook Pro). For what concerns the type of agent, participants were told that the virtual agent was controlled either by a computer or a person, similarly to the experiment conducted by Obhi and Hall (2011) [11]. Temporal binding was the dependent variable, and was calculated by finding the error means related to the perceived interval of time between the action performed by the agent and its outcome. The experiment had a between-subject experimental design as the sample was divided into two main groups of the same size depending on the agent condition (computer vs human). Each participant performed the same task by interacting with both interfaces (virtual reality and screen). All the conditions were counterbalanced. The experiment was conducted in the same laboratory for each subject, and in a similar time of the day, in order to control environmental influences and tiredness effects.

2.3. Apparatus and Procedure

Before starting the experiment, each participant read and signed the ethical consent. Afterwards, they were told to sit on a chair and wait for instructions. After being assigned to a specific agent condition, they were provided with detailed information about the task. They had to watch a pre-recorded animation, realized on Blender and Adobe Premiere, where a virtual archer threw arrows towards a target. The video was accompanied also by sound effects: the first one lasted 150 ms and occurred when the arrow was shot (Figure 2a), whereas the second one lasted 250 ms and occurred when the arrow reached the target (Figure 2b). The two events were separated by a pseudo-random interval ranging between 200 and 1400 ms. Particularly, there were 7 possible intervals. Their sequence was random, and each interval occurred the same number of times. After each trial, a black screen with the text “Recreate the interval” appeared and lasted for 2000 ms (Figure 2c). Participants had to recreate the interval between action and auditory outcome by pressing the space bar of a keyboard twice, which was connected to a stopwatch (Figure 2d). This procedure was inspired by the technique used by Poonian et al. (2015) [12]. The experiment had two blocks, one for both interfaces. Blocks had 35 trials each divided into 4 breaks of 10 seconds. Each person had a total of 70 trials, plus 2 training trials per condition. After terminating the task with the first interface, participants had 2 minutes break before staring the second half of the experiment with the other interface. At the end of the experiment, a debriefing about the real purposes of the study followed.

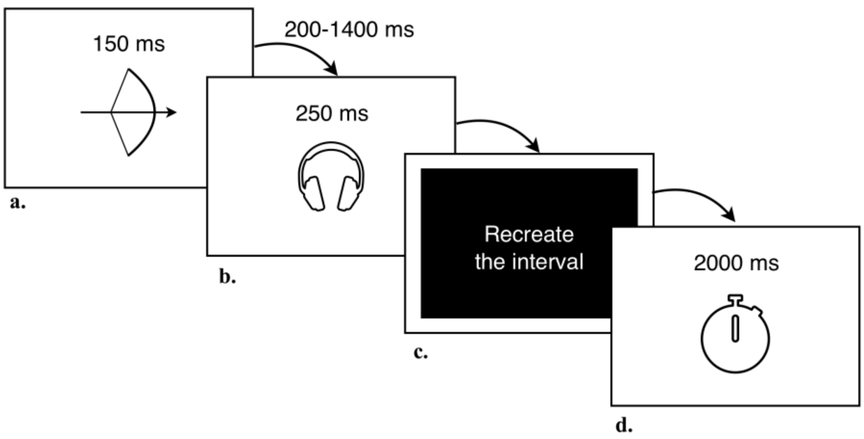

Figure 2. Graphical illustration of the task. **a.** The archer shot the arrow and a sound of 150 ms occurred. **b.** The arrow reached the target and a second sound of 250 ms occurred after a pseudorandom interval of 200-1400 ms. **c.** A black screen followed and lasted for 2000 ms. **d.** While the screen was still black, participants estimated the interval between the onset of the action and its auditory outcome by using a stopwatch (double click of the space bar).

2.4. Data analysis

For each individual trial, it was calculated the error related to the estimation of the interval of time between action and auditory outcome. To do so, the perceived interval of time was subtracted from the actual interval: positive errors indicate shorter perceived intervals, whereas negative errors represent longer estimations. After that, the outliers were found by calculating average and standard deviation (SD) for each person: all the errors which exceeded 2.5 SD from the individual mean were discarded. Subjects with a number of missing values and/or outliers greater than 10% of the total trials (7 out of 70), were eliminated from the analysis. Only two subjects had to be discarded by following these criteria. Afterwards, individual and total error means were calculated without outliers and used for the statistical analysis. First of all, some descriptive statistics were conducted to test the normality of the distribution. Afterwards, a 2-way ANOVA was ran to analyze the interaction between interface and type of agent and their effects on temporal binding. Subsequently, a Kruskal-Wallis Test was performed to test the influence of different intervals on time perception.

3. Results

Some trends emerged from a simple analysis of the general means of all conditions (Figure 3). Computer agent condition had the highest levels of temporal binding in both interface conditions (error mean in virtual reality = 72.752 ms, error mean in screen condition = 28.128 ms), whereas the error mean in human condition was very close to zero (error mean in virtual reality = 1.258 ms, error mean in the screen condition = -0.445 ms) as we can see from Figure 3a. Virtual reality was associated with less temporal binding (total error mean = 14.691 ms) and more accuracy compared to a traditional computer screen (total error mean = 36.153) as it is shown in Figure 3b. In fact, 17 people out of 26 estimated intervals more accurately in virtual reality. Finally, human condition leaded to lower temporal binding (total error mean = 50.438) compared to computer condition (total error mean = 0.406) as we can see from Figure 3c. The last finding is inconsistent with previous studies [11,20].

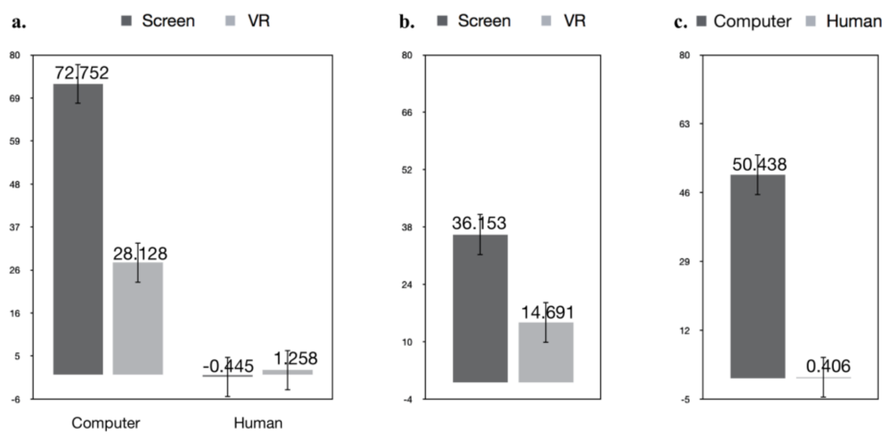

Figure 3. General error means in milliseconds across conditions. Positive values indicate higher temporal binding. **a.** Error means in computer and human conditions in both interface conditions. **b.** Total error means per interface. **c.** Total error means per type of agent.

However, these trends were not supported by the statistical analysis. In fact, a 2-way ANOVA with a significance level of .05 (two-tailed), found no significant main effects of interface and type of agent on temporal binding (F(1,48) = .301, *p* = .586 and F(1,48) = 1.636, *p* = .207, respectively), as well as their interaction (F(1,48) = . 351, *p* = .556). Thus, the test failed to reject the null hypothesis that the error means are the same in all conditions, and the alternative hypothesis could not be accepted. A reason for that, could be that the variance in the sample was high. This might be due to the insufficient size of the sample, but also to the differences in the error means related to specific intervals of time. In fact, we conducted a Kruskal-Wallis Test, to check whether the type of interval affects temporal estimation. We found a significant difference between error means in different interval conditions ($\chi^2(6) = 132.968$, $p = .000$). Particularly, people tended to perceive intervals between 200 and 600 ms as longer, and intervals between 1000 and 1400 ms as increasingly shorter (Figure 4). On the other hand, estimation of intervals of 800 ms were very accurate.

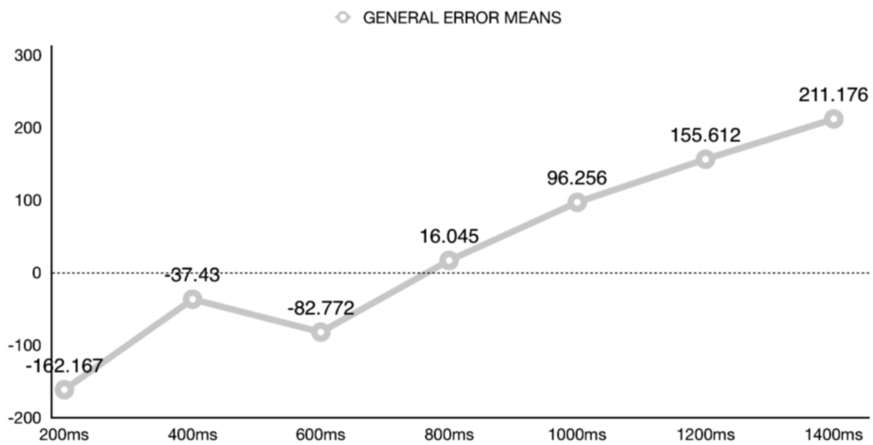

Figure 4. General error means in milliseconds in different intervals of time. The horizontal line represents an ideal performance where errors are equal to zero and the estimated time is the same as the actual interval. The grey line represents the general error means for each interval. Temporal binding corresponds to positive values.

4. Discussion

As we said, sense of agency and agency attribution are related to our ability to identify ourselves or other people as the cause of an action. These concepts are often correlated with temporal binding and rely on multiple top-down and bottom-up mechanisms and cues such as sensorimotor feedbacks, predictions, knowledge of the world and previous experience [18]. Some studies proved the existence of these multiple mechanisms underlying agency and temporal binding [4], however, not many of them focused on the influence of experimental factors. This study aimed to shed more light on these mechanisms, by focusing on the interface and the measure of agency. We also analyzed the role played by the type of agent similarly to previous studies in Human Computer Interaction [4,11]. In general, we found no significant effects of type of agent, interface and their interaction, on people's interval estimations. A reason for that, could be that the variance in the sample was high and its size was insufficient. Nevertheless, it is possible to comment some interesting trends in the data.

For what concerns the type of agent, we found inconsistent results with previous studies [11,20], as computer condition elicited higher rates of temporal binding, compared to human condition, where the interval estimation seemed to be more accurate and error means were very close to zero. A possible explanation for this unexpected result is that people thought that this specific archer was more likely to be controlled by a computer rather than a person. To reach this conclusion, they used different contextual, perceptual and inferential cues. Another explanation is that people were more accurate in the interval estimation when they thought that the archer was controlled by a person, as it is easier to predict and feel connection with human behavior. However, according to previous studies, this condition should elicit higher intentional binding levels rather than greater accuracy [11,14,15].

For what concerns the interface, we found that screen condition was associated with higher levels of temporal binding than virtual reality. This may be due to the fact that people are generally less familiar with virtual reality. However, another possible interpretation is that virtual reality reduces external world's distractions and increases accuracy of time perception and interval estimation. If future studies proved it right, virtual reality could be reasonably used in the healthcare, for example, to increase the chances of successful medical surgeries, or it could be beneficial for education in schools or distant learning. The fact that we did not find significant results in this case could be also due to the fact that the task mainly relied on auditory cues rather than visual. Thus, future studies could introduce new tasks where visual stimuli are more relevant, but in this case we suggest to avoid using measures of agency based on high visual load such as the Libet clock as they can distract from the main task [9].

In addition to the interface, the other experimental factor which we took in consideration was the measure of agency. We investigated whether temporal binding effects and time perception in general are influenced by changes in the measure of agency. Particularly, we used the interval estimation procedure and we decided to use a wider range of intervals compared to previous studies. We found that time estimations varied significantly according to specific intervals which had to be estimated. From one side, when the auditory outcome occurred 200-600 ms after the arrow was shot, intervals were perceived as longer. From the other side, when the outcome occurred 1000-1400 ms after the action, intervals were perceived as increasingly shorter. Our hypothesis is that people used both sensorimotor cues and inferential processes to elaborate context and stimuli, and estimate an optimal interval between action and outcome. As the archery scenario was realistic and familiar to people, they could extract information from their previous experience and predict the most likely interval between events: for example, they could use the distance between the archer and the target or the speed of the arrow, which did not vary over the trials. In this case, 800 ms seemed to be the most plausible interval, and the more intervals stepped away from the optimal value, the more errors tended to linearly increase.

These results not only question what we know about agency, temporal binding, and interval estimations, but are also fundamental to develop and improve existing computational models of time perception. A type of models which is more suitable in this case are cognitive architectures such as ACT-R, where different aspects of cognition are integrated and interact between each other in order to understand human functioning [1]. In general, in these models memory is not explicitly taken in consideration, except for some models which focus on working memory [1]. For example, in the Pool Model [19] there is a “temporal reference memory” which stores representations of each new interval of time based on a “pool of recent experiences”. When the actual interval corresponds to a “learned standard”, the system can recognize it [19]. Intervals can be learnt and compared thanks to a pacemaker-accumulator system which detects pulses and attributes a certain number of pulses to each interval [191]. The model takes in consideration that our estimations are always approximate by adding noise to every pulse. The equation which represents these processes (Figure 5) says that when an interval is recalled at a certain time, its value depends on the time when it has been created and on the match with the actual interval (mismatch-penalty = 0 when there is match, negative value when there is mismatch) [19].

$$A(t) = \log(t - t_{\mathrm{creation}})^{-d} + \mathrm{mismatchpenalty}$$

Figure 5. Equation of the Pool Model elaborated by Taatgen and van Rijn (2011) [19].

However, this model is mainly focused on recently-formed memories and how they degrade over the time. Our suggestion for future studies is to implement a similar model which takes in consideration also long-term and non-declarative memory. In more detail, the model would need a pool of long-lasting memories which represents our knowledge of the world. Each memory of an interval is associated also to a memory of a specific context. In this way, every time the system recognizes a certain context and detects a specific interval, it can estimate its plausibility and likelihood by comparing it with a related memory.

In conclusion, this study aimed to investigate the influence of some experimental factors such as interface and measure of agency on temporal binding. We found that both traditional screen and virtual reality elicited temporal binding but there were no significant differences, even though screen seemed to bring higher rates of temporal binding and virtual reality seemed to elicit higher accuracy. For what concerns the measure of agency, we found significant differences in the estimation of intervals depending on the intervals which had to be estimated. Particularly, we think that in each context and circumstance, people tend to predict an optimal interval between specific actions and outcomes, by using sensorimotor cues, contextual information and inferential processes. The farther actual intervals are from the optimal one, the higher estimation errors are. From one side, these results questioned what we know about agency and proved that methods such as interval estimation procedure are not always robust [3]. Thus, we hope this study can prompt researchers to be more aware of their experimental choices and careful when they interpret their results. From the other side, these findings shed more light on the mechanisms of time perception and hopefully will be beneficial for future studies for the implementation of new computational models.

References

1. Anamalamudi, K., Surampudi, B. R., Maganti, M.: Computational Models of Time Perception. In: 1st International Conference on Networks and Soft Computing, pp. 413-417. IEEE, New Jersey (2014).
2. Buehner, M. J.: Understanding the Past, Predicting the Future: Causation, Not Intentional Action, Is the Root of Temporal Binding. Psychological Science, 23(12), 1490-1497 (2012).
3. Coyle, D., Moore, J., Kristensson, P. O., Fletcher, P. C., Blackwell, A. F.: I did that! Measuring Users' Experience of Agency in their own Actions. In: CHI '12 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp 2025-2034. ACM, New York (2012).
4. Goldberg, M.: The sense of agency: underlying neurocognitive mechanisms and its attribution to human and non-human co-actors. Humboldt University of Berlin, Berlin (2018).
5. Haggard, P., Tsakiris, M.: The experience of agency: feelings, judgments, and responsibility. Current Directions in Psychological Science, 18(4), 242-246 (2009).
6. Humphreys, G. R., Buehner, M. J.: Magnitude Estimation Reveals Temporal Binding at Super-Second Intervals. Journal of Experimental Psychology: Human Perception and Performance, 35(5), 1542-1549 (2009).
7. Limerick, H., Coyle, D., Moore, J.W.: The experience of agency in Human-Computer Interactions: A Review. Frontiers in Human Neuroscience, 8, 643 (2014).
8. Limerick, H., Moore, J.W., Coyle, D.: Symbiotic interaction and the experience of agency. In: Blankertz, B. et al. (eds.). Symbiotic Interaction. 4th International Workshop, Symbiotic 2015, Berlin, Germany, LNCS, vol. 9359, pp. 99-104. Springer, Cham (2015).

9. Martinez, C., Maggioni, E., Hornbæk, K., Obrist, M., Subramanian, S.: Beyond the Libet Clock: Modality Variants for Agency Measurements. In: CHI ’18 Conference on Human Factors in Computing Systems, pp 1-14. ACM, New York (2018).
10. Moore, J. W., Teufel, C., Subramaniam, N., Davis, G., Fletcher, P. C.: Attribution of Intentional Causation Influences the Perception of Observed Movements: Behavioral Evidence and Neural Correlates. Frontiers in Psychology, 4 (23) (2013).
11. Obhi, S. S., Hall, P.: Sense of agency in joint action: influence of human and computer co- actors. Experimental Brain Research, 21(1), 663-670 (2011).
12. Poonian, S. K., McFadyen, J., Ogden, J., Cunnington, R.: Implicit agency in observed actions: evidence for N1 suppression of tones caused by self-made and observed actions. Journal of Cognitive Neuroscience, 27(4), 752-764 (2015).
13. Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L.: Premotor cortex and the recognition of motor actions. Brain Research. Cognitive Brain Research, 3(2), 131-141 (1996).
14. Sahaï, A., Pacherie, E., Grynszpan, O., Berberian, B.: Co-representation of human-generated actions vs. machine-generated actions: impact on our sense of we-agency? In: 26th IEEE International Symposium on Robot and Human Interactive Communication, pp. 341-345. IEEE, New Jersey (2017).
15. Sahaï, A., Pacherie, E., Grynszpan, O., Berberian, B: Predictive mechanisms are not involved the same way during human-human vs. human-machine interactions: a review. Frontiers in Neurorobotics 11(52), 1-18 (2017).
16. Synofzik, M., Vosgerau, G., Newen, A.: Beyond the comparator model: a multifactorial two-step account of agency. Consciousness and Cognition, 17(1), 219-39 (2007).
17. Synofzik, M., Vosgerau, G., Newen, A.: I move, therefore I am: a new theoretical framework to investigate agency and ownership. Consciousness and Cognition, 17(2), 411-424 (2008).
18. Synofzik, M., Vosgerau, G., Voss, M.: The experience of agency: an interplay between prediction and postdiction. Frontiers in Psychology, 4(127), (2013).
19. Taatgen, N., van Rijn, H.: Traces of times past: Representations of temporal intervals in memory. Memory & Cognition, 39 (8), 1546–1560 (2011).
20. Wohlschläger, A., Haggard, P., Gesierich, B., Prinz, W.: The Perceived Onset Time of Self- and Other-Generated Actions. Psychological Science. 14(6), 586-591 (2003).