Computational cognitive modelling of action awareness: prior and retrospective

Dilhan J. Thilakarathne · Jan Treur

Abstract This paper presents a computational cognitive model for action awareness focusing on action preparation and performance by considering its cognitive effects and affects from both prior and retrospective form relative to the action execution. How action selection and execution contribute to the awareness or vice versa is a research question, and from the findings of brain imaging and recording techniques more information has become available on this. Some evidence leads to a hypothesis that awareness of action selection is not directly causing the action execution (or behaviour) but comes afterwards as an effect of unconscious processes of action preparation. In contrast, another hypothesis claims that both predictive and inferential processes related to the action preparation and execution may contribute to the conscious awareness of the action, and furthermore, this awareness of an action is a dynamic combination of both prior awareness (through predictive motor control processes) and retrospective awareness (through inferential sense-making processes) relative to the action execution. The proposed model integrates the findings of both conscious and unconscious explanations for both action awareness and ownership and acts as a generic computational cognitive model to explain agent behaviour through the interplay between conscious and unconscious processes. Validation of the proposed model is achieved through simulations on suitable scenarios which are covered with actions that are prepared without being conscious at any point in time, and also with the actions that agent develops prior awareness and/or retrospective awareness. Having selected an interrelated set of scenarios, a systematic approach is used to find a suitable but generic parameter value set which is used throughout all the simulations that highlights the strength of the design of this cognitive model.

Keywords Awareness · Prior · Retrospective · Cognitive modelling

1 Introduction

Humans intuitively feel that their behaviour is an effect of their conscious decisions for certain actions aiming for desired outcomes [1]. However, what exactly is consciousness is a well-known question among many scientists in many disciplines (see [2, 3]); for example, is it just a process in the brain, and if so, how is it composed? With the developments in brain imaging and recording techniques, more and more detailed information on various brain processes becomes available, including the conscious awareness of actions. One of the leading hypotheses for action awareness is that humans may prepare for and perform actions without being conscious of these preparation and execution processes. More specifically, the feeling of intention for an action is not causing the behaviour but comes after the action preparation and just before the action execution time [4–8]. It has been found that for certain types of actions, the decision to perform it is already made at least hundreds of milliseconds (and even up to 10 s) before any awareness state occurs [4, 5, 8]. The brain

D. J. Thilakarathne (✉) · J. Treur
Agent Systems Research Group, Department of Computer Science, VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands
e-mail: d.j.thilakarathne@vu.nl

J. Treur
e-mail: j.treur@vu.nl

predicts the outcome of a decision even before the decision reaches awareness and humans' (illusionary) consciousness seems like an after-effect of a set of unconscious cognitive processes leading to the action [1, 4–8].

The human brain is a complex, intricate, adaptive, dynamic system; it is difficult to unravel it and comprehend its mechanisms (cf. [9–14]). Therefore, given the complexity and contradictions observed in different experiments, alternative hypotheses are also proposed on action awareness (cf. [15–19]). In particular, it is interesting to analyse how an individual is acquiring skills related with a task that he/she is not familiar before (i.e. before something is becoming a habitual task) and how awareness contributes in such a situation [2]. For example, someone who is in the mid age but has no prior experience of cycling, it is an interesting phenomenon to find the experiences gone through in the learning process to develop this skill. Most probably at the beginning of this task, it may be really challenging and much attention and awareness are required to perform the task. As a result of that he/she may not be able to change the (visual) focus at all to other environmental cues in order to maintain the balance, keep the ride straight, take turns safely, and pedal up a sloop. Most probably combining what is predicted and what is observed, thereby minimising the error (prediction vs. actual) is the basis for the collected experience in the learning period. Furthermore, how the awareness affects the neural plasticity is an interesting phenomenon too. Although in the learning period, a high level of awareness on the action is important, once it becomes a habitual task such awareness may practically become absent: most probably he/she may not pay any attention at all on pedalling, balancing and keeping straight. Therefore, in daily life experiences predictive and inferential processes for action awareness are important. Nevertheless, when a person learned how to ride a bicycle properly, still he/she may not able to have a small snack or holding an umbrella in a rainy time while cycling (though these can also become habitual tasks for some people). Therefore, just slightly deviating from a habitual task seems to require awareness or new experience.

Moore and Haggard [20] have investigated how predictive or inferential processes of action execution play a role in conscious awareness of an action. They have proposed that awareness of an action is a dynamic combination of both prior awareness (i.e. awareness of the action prediction) and retrospective awareness (i.e. the awareness of the effects of an action), through predictive motor control and inferential sense-making relative to the action execution, respectively (cf. [20, 21]). When a prior awareness state occurs, he/she may become aware of going to perform the action. Having such a prior awareness state still may leave open whether the agent is able to consciously decide to perform or not to perform the action (cf.

[1, 4–8]). For example, is still some form of vetoing of the action possible? In principle, the awareness state may play the role of generating a kind of green light for execution of the action. However, equally well the prior awareness state may just play the role of a warning for the agent to be prepared that the action will happen (anyway). As stated in [5, 8] it has been found that for certain types of actions, the decision to perform it is already made at least hundreds of milliseconds (and even up to 10 s) before any awareness state occurs. These findings may suggest that prior awareness often will have no effect on the decision. But this may strongly depend on the type of action. For example, it will be difficult to believe that the action of buying a car or a house remains unconscious and may not be amendable to vetoing based on awareness states (the Monty Hall Problem is also another good situation for this concern). An awareness state can also develop in retrospect, after the action was performed and this will answer the question: 'what have I done?'. Such a retrospective awareness state often relates to acknowledging others from and taking responsibility for having performed the action. It may also play an important role in learning as mentioned in the previous example (i.e. by evaluating the obtained effect in a conscious manner leads to improvement of the performance of the action selection).

This paper extends the work published in [22] by refining the neurologically inspired agent model with more realistic simulation results, new scenarios and a detailed formal specification of the model, together with a more sophisticated parameter estimation methodology. The selected scenarios include a reasonable spectrum of situations in which (a) actions are prepared without being conscious at any point in time, (b) the agent develops prior awareness or retrospective awareness or both. An example of a schizophrenic patient and of a patent in an early stage that may lead to a depression situation are also included. As research questions, this paper mainly contributes to

(1) How does the internal prediction process shape or contribute to the (prior) awareness of the action?
(2) How does the inferential sense-making shape or contribute to the (retrospective) awareness of the action execution?
(3) How does the awareness contribute to action execution?
(4) What is the relation and interplay between conscious and unconscious action formation through action ownership and relevant awareness states?

In addition to awareness states, ownership states for an action are also considered in this paper. They are mainly used as important states in the unconscious action formation process. The specific role of the ownership states (in prior and retrospective form) has been separately discussed

in [23]; such a more detailed overview is not included in this paper. The structure of the paper organised with a conceptual basis that includes evidences from cognitive neuroscience which followed with a model description in which a detail explanation of the model together with its mathematical basis and formal specification will be presented. To validate the workings of the proposed model, eight scenarios are simulated through a unique parameter value set which was estimated using a systematic approach. Finally, a discussion will be presented highlighting the usefulness of a model in this nature and future works.

2 Action awareness viewed neurologically, psychologically and behaviourally

In neurological, psychological and behavioural literature, the notions of awareness and ownership of an action have received much attention. Action ownership is a useful concept which is mainly important to differentiate in how far a person attributes an action to him or herself, or to another person (see [23]). Although in many cases, the feeling that you get when you perform something or another person is performing the same action may be similar, it is clearly possible to identify whether the action belongs to you or to someone else. More importantly, the information about another person's behaviour influences your self evaluation and vice versa, which makes humans social beings (cf. [24, 25]). After the discovery of mirror neurons, such social phenomena including empathy, imitation, and coordination in a social context can be explained more scientifically as a cognitive process [26–28]. Mirror neurons have been mainly identified in two cortical areas: the posterior part of the inferior frontal cortex and the anterior part of the inferior parietal lobule [28]. They have shown strong correlations not only with specific movements, but also with specific goals (or goal directed actions: e.g. reaching for and grasping an object). From the development perspective of human cognition on self and other representations, their interconnection and how those relate to the cognitive processing were highlighted:

Over the first several years of life, children acquire knowledge of both objective and subjective aspects of self and others. By 18–24 months of age infants can recognize their own mirror image, a capacity that has been linked to the emergence of self-conscious emotions (e.g. embarrassment [...]). During the preschool years, children simultaneously develop the capacity to represent their own and others' mental states [...]. This development entails the ability to recognize when self and other perspectives and experiences are shared and thus congruent, and under

which circumstances they differ from one another. Interestingly, the development of mental state understanding is functionally related to executive functions [...], suggesting that the prefrontal cortex is implicated in self/other cognitive representations. Indeed, neuroimaging data suggest that theory of mind tasks and executive function tasks share overlapping areas of activation in the medial prefrontal cortex ([24], pp. 527–528)

This separation of self and other is contributing to the ability to recognise ownership of your own action. Research has shown evidence that action prediction (based on sensory information) leads to an action execution with ownership, while when there are problems with action prediction that leads to abnormal states of ownership of that action. For example, when you are tickled by someone, as you have not predicted the action you will experience various sensations due to this sudden action, and the ownership of that action may not be with you (though you do have the body ownership in this situation) (cf. [29]). Chaminade and co-workers [25] have highlighted this:

The neural underpinnings of internal models for motor control have been investigated with human non-invasive neuroimaging techniques (for review see [30]). Motor commands that are used by forward models to suppress sensory signals are believed to originate upstream from the primary motor cortex [31], though they may also involve premotor areas in the posterior inferior frontal gyrus [32]. Actual sensory feedback is used to compute prediction errors for model evaluation and update. When we are tickled by another person [33] the sensory consequences of its actions are unpredictable, and the lack of predictability leads to a high prediction error associated with increased activity in the secondary somatosensory cortex. This area, located bilaterally in the parietal opercula [34], plays a key role in sensorimotor integration [35], and has been involved in the assessment of action ownership ([25], p. 2, [36])

The nature of human actions varies from direct responses on stimuli to actions that take longer periods to process and react. Here the first types of action are often labelled as automatic or unconscious and the other types as more conscious or intentional [37]. In contrast to action ownership, action awareness is a conscious state. Patients with the anarchic hand syndrome (AHS) [patients with frontal lobe and callosal damage (cf. [38])] always have some form of ownership and awareness of their action but are not able to control the action [39]. For example, in a cafe just seeing a cup of coffee of an unknown person, for

an AHS patient might be sufficient to reach and grasp it due to the automatic activation of action plans on this habitual task. In normal context this has been trained as a habitual task, but with the (prior) awareness of people who know when he or she should do this. A few more of such AHS examples are grabbing a doorknob or scribbling with a pencil or combing one's hair [38]. Furthermore, persons suffering from schizophrenia may easily attribute self-generated actions to (real or imaginary) other persons (see [23]). Furthermore, it has been noted that the problem with AHS patients is to control their action, while with a schizophrenic patient it is a problem with awareness of the action (an AHS patient tries to prevent abnormal behaviour of the alien hand by the good hand after it executed the action) [38]. Through learning with intention and awareness, people have pre-stored actions per stimulus and later without the intention or the awareness the brain will automatically evoke the relevant action which was habitually associated [40]. The frequency and recency of a learned habitual task seem to relate to its probability of getting selected.

The research on action awareness is a challenging task and it is assumed to be that individuals are aware only of the tip of the action iceberg; much further research is necessary to explore and refine the body of knowledge on this (cf. [2, 3, 8, 11, 15, 21]). Nevertheless, there are interesting research findings on this. Empirical evidence collected through an experiment setup proposed by Benjamin Libet and his colleagues [8] has challenged the traditional view of human will and has shown that the brain initiates voluntary movements before we are aware of having decided to move. From a cognitive neuroscience perspective, human actions are mainly a result of signals getting to motor neurons (motoneuron) in the spinal cord mainly via the primary motor cortex and some of its neighbour areas [e.g. pre motor cortex, supplementary motor cortex (SMA)]. Early activation of the primary motor cortex before the agent gets the conscious intention to move (or to act) is called readiness potential and this begins hundreds of milliseconds or even up to 10 s before any awareness state occurs [4, 5, 8]. Therefore, it was proposed that conscious will is an illusion and it is too slow to initiate an action, but action formation is due to an unconscious causal chain of processes and just before the action execution, we will develop the awareness of the action (not as the cause of the action but as an effect of unconscious processes). John-Dylan Haynes has further improved the Libet experiment setup to advance beyond the shortcomings of the experiment (see [5]) and with his findings again the importance of exploring the tightness of the link between unconscious predictive brain processes and subsequent decisions from a conscious perspective is highlighted:

An important point that needs to be discussed is to what degree the finding of choice-predictive information supports any causal relationship between brain activity and the conscious will. Such causal links have been demonstrated previously by direct cortical stimulation over parietal and frontal cortex.[37,80] However, it is unclear if the early predictive signals are also causally involved in the decision. As for the criterion of temporal precedence, there should be no doubt that our data finally demonstrate that brain activity can predict a decision long before it enters awareness. A different point is the criterion of constant connection. A constant connection would require that the decision could be predicted with 100 % accuracy from prior brain activity. Libet's original experiments were based on averages, so no statistical assessment can be made about the accuracy with which decisions can be predicted. Our prediction of decisions from brain activity is statistically reliable, but far from perfect. The predictive accuracy of around 60 % (which is significant, but only 10 % above chance) can be improved if the decoding is tailored to each subject. However, even under optimal conditions, this is far from 100 % for several reasons… … …. Importantly, a different interpretation could be that the inaccuracy simply reflects the fact that the early neural processes might only be partially predictive of the outcome of the decision. In this view, even full knowledge of the state of activity of populations of neurons in FPC and in the precuneus might not permit the full prediction of a decision. In that case, the signals have the form of a biasing signal that influences the decision to a degree, but additional influences at later time points might still play a role in shaping the decision. The fact that decoding after the decision from motor cortex can be achieved with higher accuracy might point toward the fact that neural signals in BA10 and in PC are not fully predictive in principle. However, the exact topology of clustering of calls with similar tuning preferences in BA10/PC is, to date, unknown, and thus might turn out to be less suitable for fMRI decoding than in motor cortex ([5], pp. 16–17)

With the concerns highlighted in the above quote (for more criticisms on this hypothesis see [41]), though the awareness state emerges just before the action execution, it is not yet clear whether there is not at all an impact on action execution from this subjective awareness. One of the issues that have turned out to play an important role both in the execution decisions for an action, and in its attribution, is the prediction of the (expected) effects of the action, based on internal simulation starting from the preparation

of the action [42, 43]. If these predicted effects are satisfactory, this may entail a 'go' decision for the execution of the action, thus exerting control over action execution. In contrast, less satisfactory predicted effects may lead to a 'no go' decision (cf. [44–46]). Predicted action effects also play an important role in attribution of the action to an agent after it has been performed. In neurological research, it has been found that poor predictive capabilities are a basis for false attributions of actions, for example, for patients suffering from schizophrenia [38, 47, 48]. In addition to the predictive effects, the sensation of the actual effect (after executing the action) also has been noted as important in action formation research [15, 17, 21, 37, 43]. In literature, it has been reported that the predicted sensory effect and the sensed actual effect are integrated with each other as a basis for proper attribution of the action [20, 47, 48]. Another element, put forward in [20], is the distinction between action awareness based on prediction (prior to execution), and action awareness based on an inference after execution of the action (in retrospect):

> Our results suggest that both predictive and inferential processes contribute to the conscious awareness of operant action. The relative contribution of each of these processes seems to be context dependent. When we can predict the consequences of our actions, as in a high action-effect contingency block, the awareness of action reflects these predictions. This would provide us with a predictive sense of our own agency. In addition, our results show clear evidence that inferential processes also influence the conscious awareness of operant action … … …. The interaction between predictive and inferential processes is of particular interest … … …. The time course over which information about action is built up may be an important clue to this interaction … … …. Sensory feedback provides more precise evidence about actions and their effects. This evidence becomes available only after a short sensory delay, but can then be transferred to memory. Thus, reliable and enduring sensory evidence replaces short-lived predictive estimates. We suggest that awareness of action therefore switches from a predictive to an inferential source as the action itself occurs, and as sensory information becomes available."([20], pp. 142–143)

With these evidences, they have suggested that awareness of an action is a dynamic combination of both prior awareness (i.e. awareness of the action-effect prediction) and retrospective awareness (i.e. the awareness of the effects of an action) through predictive motor control and inferential sense-making relative to the action execution, respectively [15, 21, 37, 43] (cf. [19]). Furthermore,

Haggard and co-workers presented a new phenomenon called intentional binding: when a voluntary action produces (with the awareness and intention) the temporal (subjective) gap between the action and its perceived sensory outcome is less when the awareness is pre-existing but it is high when the awareness does not involve this [15]. This phenomenon has been argued to be an effect of either prior awareness or retrospective awareness with different experiment setups. To investigate the relation with prior awareness, transcranial magnetic stimulation (TMS) was randomly applied over the motor cortex and the entailed disruption of awareness observed (here through the intention in this setup). A significantly weakened intentional binding has been observed; therefore, it may be useful to highlight the necessity of prior awareness (see [19]). Similarly, there are experiments to analyse the influence from retrospective awareness selecting some tasks where prediction of the action outcome is difficult (or unpredictable) but there is a 'tone' after the action execution [20] and it has been observed that retrospective processes play a role when prior predictive processes are absent (or when prediction was minimal). In addition to the mentioned roles in prior and retrospective effects of intentional binding, there are some evidences for its neural basis also:

> Moore and colleagues investigated the contribution of two specific target sites: the pre-supplementary motor area (pre-SMA) and primary motor cortex (M1). The pre-SMA is involved in higher-order cognitive aspects of self-generated action [49] and with the conscious experience of intending to act [50]. In this sense it is likely to support predictive contributions to intentional binding. On the other hand, M1 processes signals that are involved in actual motor execution, signals that the authors suggest are required to support inferences of agency … … …. It was found that only stimulation of pre-SMA led to a significant reduction in intentional binding. Stimulation of M1 marginally reduced intentional binding, but this effect was not significant. The authors therefore concluded that pre-SMA is likely to play a key role in intentional binding. ([19], pp. 5)

Inhibition and suppressive mechanisms may also be as important as the excitation mechanisms in cognitive control [though some different viewpoints are also put forward (see [51])]. By Gamma-aminobutyric acid (GABA), neurons are performing inhibition at synaptic, circuit and systems levels (cf. [51]). Furthermore, various inhibition types in neuroscience and psychology have been discussed in [51]. Inhibition activates in automatic (e.g. lateral inhibition: if a particular representation accumulates more evidences, that will suppress its fellow representations) and voluntary (e.g. suppression of an irrelevant response,

stimulus or memory; in intentionally) manners. Another peculiar aspect that has been observed is that within the process of co-occurrence of predicted effects and sensed actual effects, the predicted effect suppresses the sensed actual effect [29, 52, 53]. Moreover, it has been put forward that the predicted effect and the sensed actual effect are not simply compared or matched, as claimed in the so-called 'comparator model' in earlier literature such as [38, 42, 54], but in fact are added to each other in some integration process [20, 47, 48].

Though these evidences are facilitating an adequate level of information to comprehend a theoretical cognitive system in the form of a model, it is further required to confirm these findings; this may have different variants to be explored in future research.

3 Description of the cognitive computational model

Having discussed the evidence on awareness (a person's subjective experience) and ownership (in how far does a person attribute an action to him or herself or to another person) in Sect. 2, this section presents a computational agent model. This model will be used in agent-driven applications where the awareness is paramount (or necessary) for decision-making and justifications of actions through communication. More specifically, this model can provide interesting and important input for problem domains concerning performing or learning specific healthy behaviours or lifestyles. In such domains, having an idea about the extent of the awareness of decisions concerning health or

lifestyle is important and the model may provide the fundamentals for applications in these domains (with further refinements and customization where needed). Furthermore, this model may be useful in medical domains where people can analyse and compare the phenomenal effects of certain scenarios through various simulations and further to compare and contrast different hypothesis in theoretical manner to conduct more complex experiments. Also computational simulations have become a promising approach to analyse the emergence in complex systems (e.g. in the aviation domain, or in social science) and a model like this may provide more realistic results especially when human cognitive aspects are necessary in such simulations (e.g. simulating situation awareness in air traffic control, including the human cognitive aspects).

An overview of the postulated cognitive agent model is presented in Fig. 1 and its abbreviation details can be found in Table 1. The model is a refined version of a previous model presented in [22] but improving the action preparation process and simulation results. In this model, awareness states are taken specific for a given action a, effect b, context c and stimulus s. When the context c is self, an awareness state for a, b, c and s indicates self-attribution awareness, whereas for context c, an observed agent B, it indicates awareness of attribution of the action to B. Specific attention is given to 'self' than 'other' in simulations of this paper (for 'other' see [22, 23]). Furthermore, causal relationships in the model are based on the neurological literature presented in the Sect. 2; they do not take specific neurons into consideration but use more abstracted cognitive or mental states for the design of the

Fig. 1 Overview of the computational cognitive agent model. Here an *arrow* → represents a direct activation to state B from state A, an *arrow* → represents a direct suppression to state B from state A, an *arrow* ⇢ represents a suppression to all the complements of 'ith' state on B_i from state A_i (where 'i' presents an instance of a particular state), and → represents a direct supression to all parellel forms of that state
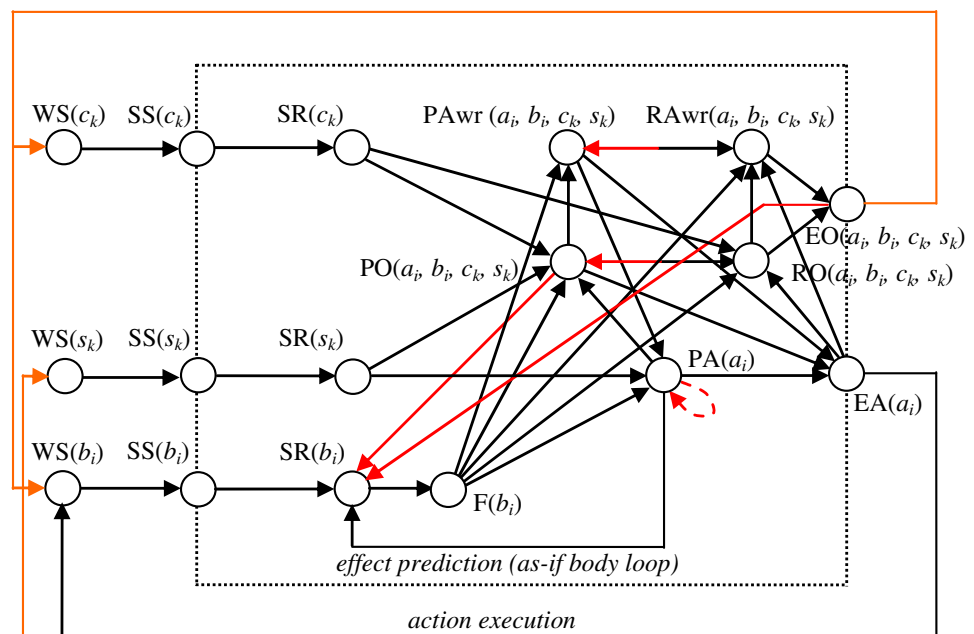
Table 1 Nomenclature for Fig. 1

WS(W)	World state W (W can be either: context c_k, stimulus s_k, or effect b_i)
SS(W)	Sensor state for W (W can be either: context c_k, stimulus s_k, or effect b_i)
SR(W)	Sensory representation of W (W can be either: context c_k, stimulus s_k, or effect b_i)
PA(a)	Preparation for action a_i
F(b_i)	Feeling for action a_i after as-if loop or action execution
EA(a_i)	Execution of action a_i
PO(a_i, b_i, c_k, s_k)	Prior ownership state for action a_i with b_i, c_k, and s_k
RO(a_i, b_i, c_k, s_k)	Retrospective ownership state for a_i with b_i, c_k, and s_k
PAwr(a_i, b_i, c_k, s_k)	Prior-awareness state for action a_i with b_i, c_k, and s_k
RAwr(a_i, b_i, c_k, s_k)	Retrospective awareness state for action a_i with b_i, c_k, and s_k
EO(a_i, b_i, c_k, s_k)	Communication of ownership and awareness of a_i with b_i, c_k, and s_k

model (through an interlevel relation between the neurological level and the cognitive/affective mental modelling level). The model uses three world states (WS) as inputs for:

stimulus s : WS(s)

context c : WS(c)

effect b : WS(b)

The stimulus s represents any internal (bodily, e.g. self-generated facial expression) or external change that may lead to an action execution. Context c represents additional information perceived to improve the process of action selection. The context c can be differentiated as 'self' and 'other' (self-other distinction). Effects of mirroring can be modelled when c is 'other' (see [22, 23]). The effect b represents the effects of the execution of an action a.

The input world states WS(s), WS(c) and WS(b) lead to sensor states SS(s), SS(c) and SS(b), and subsequently to sensory representation states SR(s), SR(c) and SR(b), respectively. The unconscious causality of action formation has been modelled as explained in [55]: by combining Damasio's as-if body loop (see [44–46]) and James's body loop (see [56]) hypotheses. The body loop has been mapped in this model by the following causal relationships:

$$\text{WS}(s) \rightarrow \text{SS}(s) \rightarrow \text{SR}(s) \rightarrow \text{PA}(a) \rightarrow \text{EA}(a) \rightarrow \text{WS}(b)$$
$$\rightarrow \text{SS}(b) \rightarrow \text{SR}(b) \rightarrow \text{F}(b)$$

Damasio extended the body loop concept and argued that the cognitive process of action selection is due to an effect of an internal simulation process prior to the execution of an action [44–46]. The brain will evaluate the effect of each relevant action option [i.e. PA(a_i)] by comparing the feelings associated to each individual valuated effects (without actually executing them through the body loop). The simulated option that has the strongest valuated feeling performs as a GO signal through the body loop and

else are NO–GO options. The as-if body loop consists of:

$$sensory\ representation \rightarrow preparation\ for\ bodily\ changes$$
$$\rightarrow felt\ emotion$$

In this model, this is represented by the as-if body loop as follows:

$$\text{PA}(a) \rightarrow \text{SR}(b) \rightarrow \text{F}(b)$$

The as-if body loop and the body loop demonstrate the working of predicted sensory effects and sensed actual effects, respectively, as highlighted in the Sect. 2. These processes are mainly considered to be unconscious processes involving multiple options for responses evaluated in parallel, to determine an adequate response associated to a stimulus (cf. [40]). Through this parallel internal action simulation mechanism, the agent will not select a random option but the one which has the strongest valuated feeling. Therefore, depending on the weight values attached to each option at that particular moment, the model will show different behaviours in simulations. This purely unconscious mechanism may be interrupted by the effects of awareness (which will be explained later) to select something different by adding some bias to the mentioned process. Being a cyclic process, the effects of an injected bias may have the ability to compete with other options to finally provide a GO signal.

In Fig. 1, state labels are attached with subscript letters k and i, which indicate, for example, the kth instance for a stimulus s (e.g. WS(s_k)) for a given s_k stimulus and the ith option for an action a (e.g. PA(a_i)). Therefore, through this model, it is possible to have multiple action options either through a single stimulus or from multiple stimuli, depending on the specific model instance.

Each PA(a_i) state is affected by its associated feeling through the as-if body loop. Moreover, each PA(a_i) state suppresses its complementary options PA(a_j) for $j \neq i$ (as shown in dotted looped red arrow in Fig. 1) proportional to the accumulated strength of that option. This behaviour is

in line with the explanation for the lateral inhibition in [51] and will contribute to further strengthen the action selection process. Therefore, naturally the strongest internally satisfied option (which is exceeding a threshold value) will become selected as a result of the unconscious action selection process as explained earlier. The feeling state in this model can be either a positive feeling or a negative feeling. A given stimulus s_k may trigger multiple preparation options in parallel and those might have different associated feelings, also in parallel (note that when a feeling state's activation level is '0' it is assumed to be a case of no feeling).

Prior ownership states have been integrated with the above-mentioned processes and mainly they get affected from sensory representation states $SR(s_k)$ and $SR(c_k)$, action preparation state $PA(a_i)$ and feeling states $F(b_i)$. Also the $PO(a_i, b_i, c_k, s_k)$ states affect prior awareness states $PAwr(a_i, b_i, c_k, s_k)$, retrospective ownership states $RO(a_i, b_i, c_k, s_k)$, action execution states $EA(a_i)$ and sensory representation states $SR(b_i)$ of effects b_i. Having a direct link between $SR(s_k)$ and $PO(a_i, b_i, c_k, s_k)$ facilitates the embedding of salient features of the input to the ownership and therefore the agent will be able to relate the input and output [together with $RO(a_i, b_i, c_k, s_k)$]. Furthermore, the link from $SR(c_k)$ to $PO(a_i, b_i, c_k, s_k)$ facilitates the necessary behaviour of mirror neurons when c is other (for more details see [23]). The state $PO(a_i, b_i, c_k, s_k)$ has a suppressive effect on $SR(b_i)$; this provides the mechanism by which the predicted effect suppresses the sensed actual effect (see [29, 52, 53]). Similar to prior ownership, once an action is executed retrospective ownership will develop. The retrospective ownership state $RO(a_i, b_i, c_k, s_k)$ is affected by the prior ownership state $PO(a_i, b_i, c_k, s_k)$, $SR(c_k)$, $F(b_i)$ and $EA(a_i)$. Furthermore, $RO(a_i, b_i, c_k, s_k)$ activation has effects on the states $RAwr(a_i, b_i, c_k, s_k)$, $PO(a_i, b_i, c_k, s_k)$ and $EO(a_i, b_i, c_k, s_k)$. As RO is affected by $EA(a_i)$ and $F(b_i)$, this provides the cognitive behaviour of retrospective effects as differentiated from prior behaviour. For more details on ownership states of this model, see [23]. Once $RO(a_i, b_i, c_k, s_k)$ developed it has a suppressive effect on $PO(a_i, b_i, c_k, s_k)$ and through this also it is contributing to the cognitive shift from predictive to inferential.

For each ownership state, an associated awareness state may (or may not) emerge. Awareness states play a higher order cognitive role. The direct links from ownership and feeling to awareness state realise bottom–up activation. Conversely, the effects of awareness states on other states realise top–down activation, which is considered to be a conscious or intended process. Therefore in the presented model, $PAwr(a_i, b_i, c_k, s_k)$ is only affected by $PO(a_i, b_i, c_k, s_k)$ and $F(b_i)$. This is useful to model the idea of Benjamin Libet and others: brains initiate voluntary movements before we are aware of having decided to move (also see the simulations in Sect. 4). Moreover $PAwr(a_i, b_i, c_k, s_k)$ affects $PA(a_i)$ and $EA(a_i)$ and this is reflects the idea of Haggard and co-workers: there may be an impact from this subjective awareness state on action execution. By this $PAwr(a_i, b_i, c_k, s_k)$ to $PA(a_i)$ link the agent can inject some bias to the current unconscious process through awareness. This may strengthen a weaker action option and improve the predictive feeling of that option (which may lead to getting it executed). Furthermore, in this model $PAwr(a_i, b_i, c_k, s_k)$ can also directly strengthen the action execution state. Both the prior ownership states and the prior awareness states are associated to the predictive aspects of the system. In contrast, retrospective awareness states are associated to the inferential aspects, as highlighted in Sect. 2: awareness of an action is a dynamic combination of both prior awareness and retrospective awareness through predictive motor control and inferential sense-making relative to the action execution. Once the $RAwr(a_i, b_i, c_k, s_k)$ state is activated, it has a suppressive effect on $PAwr(a_i, b_i, c_k, s_k)$ and due to this $PAwr(a_i, b_i, c_k, s_k)$ will weaken and $RAwr(a_i, b_i, c_k, s_k)$ will be dominant after the action execution. Finally, acknowledging of ownership and awareness of an action is modelled by the connection from the $RO(a_i, b_i, c_k, s_k)$ and $RAwr(a_i, b_i, c_k, s_k)$ to the $EO(a_i, b_i, c_k, s_k)$. Once the state $EO(a_i, b_i, c_k, s_k)$ has become activated, it has a suppressive effect on $SR(b_i)$ so that this will allow to stop the inferential sense-making process.

In addition to the above mentioned connections a few more suppressive connections are available, which are shown in orange arrows in Fig. 1. These connections are mainly for purposes of having an appropriate scenario. Once a stimulus s and context c are activated, the agent starts to activate the internal processes as mentioned above and once the agent performed the action, a mechanism is assumed that stops the stimuli as an action effect: the agent has performed the task and due to that environment has changed. Therefore these orange connections: $EA(a_i)$ to $WS(s_k)$, $EO(a_i, b_i, c_k, s_k)$ to $WS(c_k)$, and $EO(a_i, b_i, c_k, s_k)$ to $WS(b_i)$ have been included to stop the input stimulus s_k, input context c_k and the effect b_i of action a_i. Furthermore, having two inputs (i.e. s_1, s_2, c_1 and c_2) if only one action is executed [let's say $i = 1$: $EA(a_1)$ and $EO(a_1, b_1, c_1, s_1)$] then it will be assumed that the executed action will suppress all the inputs, for the purpose of an appropriate scenario.

The following is a brief summary of the agent's internal causality when given stimulus s_k and context c_k as inputs:

(1) *action-effect prediction* sensory representation of effect b_i is affected by preparation of an action a_i

(2) *preparation for action a_i* is affected by sensory representation of s_k, prior-awareness, feeling of

effect prediction of action a_i and the complements of current preparation for action a_i

(3) a *prior ownership state* is triggered based on preparation for action a_i, predicted effects b_i of a_i, stimulus s_k, retrospective ownership and context c_k

(4) a *prior awareness state* is activated based on feeling of the predicted effect, prior ownership and retrospective awareness

(5) *execution of action a_i* is affected by prior-awareness, prior ownership and preparation for action a_i

(6) a prior ownership state and prior awareness state exert *control over the execution* of a prepared action (go/no–go decision, vetoing)

(7) *suppression of the sensory representation of effect b_i* by both prior ownership and communication of ownership and awareness

(8) *suppression of the prior ownership state* when the retrospective ownership state is developed

(9) *suppression of the prior awareness state* when the retrospective awareness state is developed

(10) a *retrospective ownership state* is activated based on co-occurrence of predicted action effects and action effects sensed afterwards

(11) a *retrospective awareness state* is activated based on action effects sensed by execution of action a_i, retrospective ownership and prior-awareness

(12) a retrospective ownership state and retrospective awareness are internal states that also can lead to *acknowledging authorship* of the action (individually), for example, in a social context

(13) execution of an action a_i *affects the stimulus s_k* in the world

(14) communication of ownership and awareness *affects both context c_k in the world and effect b_i of action a_i in the world*

3.1 Dynamics of the model

Connections between state properties (the arrows in Fig. 1) have weights ω_k, as indicated in Table 2. In this table, a weight ω_k has a value between -1 and $+1$ and may depend on the specific context c_k, stimulus s_k, action a_i and/or effect b_i involved. By varying these connection strengths, different possibilities for the repertoire offered by the model can be realised and can be aligned with the considered scenario and behaviour. Usually weights are assumed to be nonnegative, except for the inhibiting or suppressive connections. The behaviour of the model (through simulations) depends on the values of each of these weights (together with other parameters). Determining proper values for these parameters is a non trivial task.

Table 2 Overview of the connections and their weights

from states	to state	weights	LP
$EA(a_i)$	$WS(s_k)$	ω_1	LP1
$EO(a_i, b_i, c_k, s_k)$	$WS(c_k)$	ω_2	LP2
$EA(a_i)$, $EO(a_i, b_i, c_k, s_k)$	$WS(b_i)$	ω_3, ω_4	LP3
$WS(s_k)$	$SS(s_k)$	ω_5	LP4
$WS(c_k)$	$SS(c_k)$	ω_6	LP5
$WS(b_i)$	$SS(b_i)$	ω_7	LP6
$SS(s_k)$	$SR(s_k)$	ω_8	LP7
$SS(c_k)$	$SR(c_k)$	ω_9	LP8
$PA(a_i)$, $SS(b_i)$, $PO(a_i, b_i, c_k, s_k)$, $EO(a_i, b_i, c_k, s_k)$	$SR(b_i)$	ω_{10}, ω_{11}, ω_{12}, ω_{13}	LP9
$F(b_i)$, $SR(s)$, $PAwr(a_i, b_i, c_k, s_k)$, $PA(a_j)$ $(j{\neq}i)$	$PA(a_i)$	ω_{14}, ω_{15}, ω_{16}, ω_{17}	LP10
$SR(b_i)$	$F(b_i)$	ω_{18}	LP11
$SR(s)$, $SR(c)$, $PA(a_i)$, $F(b_i)$, $RO(a_i, b_i, c_k, s_k)$	$PO(a_i, b_i, c_k, s_k)$	ω_{19}, ω_{20}, ω_{21}, ω_{22}, ω_{23}	LP12
$PO(a_i, b_i, c_k, s_k)$, $F(b_i)$, $RAwr(a_i, b_i, c_k, s_k)$	$PAwr(a_i, b_i, c_k, s_k)$	ω_{24}, ω_{25}, ω_{26}	LP13
$PA(a_i)$, $PO(a_i, b_i, c_k, s_k)$, $PAwr(a_i, b_i, c_k, s_k)$	$EA(a_i)$	ω_{27}, ω_{28}, ω_{29}	LP14
$EA(a_i)$, $F(b_i)$, $PO(a_i, b_i, c_k, s_k)$, $SR(c)$	$RO(a_i, b_i, c_k, s_k)$	ω_{30}, ω_{31}, ω_{32}, ω_{33}	LP15
$EA(a_i)$, $RO(a_i, b_i, c_k, s_k)$, $F(b_i)$, $PAwr(a_i, b_i, c_k, s_k)$	$RAwr(a_i, b_i, c_k, s_k)$	ω_{34}, ω_{35}, ω_{36}, ω_{37}	LP16
$RO(a_i, b_i, c_k, s_k)$, $RAwr(a_i, b_i, c_k, s_k)$	$EO(a_i, b_i, c_k, s_k)$	ω_{38}, ω_{39}	LP17

Here the *red colour* ω_k indicates negative weights

In this table, the column LP refers to the (temporally) Local Properties: LP1 to LP17 (cf. [57]) and that specifies the update dynamics of the activation value of the 'to state' based on the activation levels of the 'from states'. For the dynamics of each local property, a LEADSTO formalisation is used, which has been shown to be an appropriate approach to model dynamic behaviours of computational cognitive models [57]. LEADSTO is a hybrid modelling language in which a dynamic property or temporal causal relation $a \rightarrow b$ denotes that when a state property a (or conjunction thereof) occurs, then after a certain time delay, state property b will occur (see [57] for the relevance and benefits of LEADSTO in dynamic models). LEADSTO can be compared to Linear Temporal Logic, but differs in the sense that predicate logical state expressions can be used and also real numbers in them. The traces generated for LEADSTO can be seen as continuous time models satisfying the finite variability property: between any two time points there are only a finite number of state changes.

The time delay defined in LEADSTO is taken as a uniform time step Δt here. Table 3 below summarises the formalisation of LP both in LEADSTO format and in differential equation format. This is used as the formalisation of the computational form of the cognitive model described. During the processing, each state property has a strength represented by a real number between 0 and 1

through variables V (with subscripts) that run over these values. In dynamic property specifications, this is added as a last argument in the state property expressions. This representation is considered only for the LEADSTO based formalisation. Therefore, the unary predicate representation of each state in the Table 2 was extended to a binary predicate representation by including the state strength V_n [e.g. $EA(a_i)$ to $EA(a_i, V_n)$]. Furthermore, the temporal step in the original LEADSTO formalisation (see [57]) is used as a uniform time step Δt in this paper. As an example let us consider LP9 according to the LEADSTO formalisation:

LP9 sensory representation for an effect b_i state

If the preparation state for action a_i has level V_1
and the sensor state for effect b_i has level V_2
and the prior self-ownership of action a_i for b_i, c_k, and s_k has level V_3
and the communication of ownership and awareness of a_i for b_i, c_k, and s_k has level V_4
then after time duration Δt the state sensory representation for an effect b_i will have the level:
$V_5 + \gamma \, [f(\omega_{10}V_1, \omega_{11}V_2, \omega_{12}V_3, \omega_{13}V_4) - V_5 \,] \Delta t$

This expresses that after time duration Δt, the value for the sensory representation $SR(b_i)$ of effect b_i has changed from V_5 into

$$V_5 + \gamma[f(\omega_{10}V_1, \omega_{11}V_2, \omega_{12}V_3, \omega_{13}V_4) - V_5]\Delta t$$

This means that

$$\Delta SR(b_i) = \gamma[f(\omega_{10}V_1, \omega_{11}V_2, \omega_{12}V_3, \omega_{13}V_4) - V_5]\Delta t$$

or

$$\Delta SR(b_i) = \gamma[f(\omega_{10}.PA(a_i), \omega_{11}.SS(b_i), \omega_{12}.PO(a_i, b_i, c_k, s_k), \omega_{13}.EO(a_i, b_i, c_k, s_k)) - SR(b_i)]\Delta t$$

This expression in difference equation format can be rewritten into differential equation format:

$$\frac{dSR(b_i)(t)}{dt}$$
$$= \gamma \left[f\begin{pmatrix} \omega_{10}.PA(a_i)(t), \omega_{11}.SS(b_i)(t), \omega_{12}.PO(a_i, b_i, c_k, s_k)(t), \\ \omega_{13}.EO(a_i, b_i, c_k, s_k)(t) \end{pmatrix} \right.$$
$$\left. - SR(b_i)(t) \right]$$

The same formalisation but specifically for the rate of activation change can be presented in its differential form as found in Table 3 under LP9s row. Here, f is a function for which different choices can be made. The function f should be a combination function (when a given state has only a single input the identity function $f(W) = W$ is also usable though it is less configurable). For the simulations, the combination function f is based on a continuous logistic

threshold function $g(\sigma, \tau, X)$ is used as in the Eqs. (1) and (2).

$$f(\omega_{1i}y_1, \omega_{2i}y_2, \ldots) = g\left(\sigma, \tau, \sum_j \omega_{ji}y_j\right)$$

with

$$g(\sigma, \tau, X) = \left(\frac{1}{1 + e^{-\sigma(X-\tau)}} - \frac{1}{1 + e^{\sigma\tau}}\right)(1 + e^{\sigma\tau}) \quad (1)$$
when $x > 0$

$$g(\sigma, \tau, X) = 0; \quad \text{when} \quad X \le 0 \quad (2)$$

In the above equations, σ is the steepness and τ the threshold; these are configuration parameters that change the shape of the curve and its midpoint on the X-axis. Activation of a state depends on multiple other states that are directly attached to it; therefore incoming activation levels from other states are combined to some aggregated input and perform the activation according to a specification as in LP9 above, or in an alternative differential equation format, as in Eq. (3) [where g is the logistic function specified in the Eqs. (1) and (2), and y_i is the activation level of state i].

$$\frac{dy_i}{dt} = \gamma_i \left[g\left(\sigma, \tau, \sum_j \omega_{ji}y_j\right) - y_i \right] \quad (3)$$

Parameter γ is an update speed factor, indicating the speed by which an activation level is updated upon received input from other states. In this model two speed factor values are used: one for the internal states (states which are inside the dotted box in Fig. 1), and the other for the external states: $WS(s_k)$, $WS(c_k)$, $WS(b_i)$, $SS(s_k)$, $SS(c_k)$, $SS(b_i)$, $EA(a_i)$ and $EO(a_i, b_i, c_k, s_k)$. The internal states' speed factor is higher than the external states (adhering to the phenomenon that brain neurons are activating much faster than sensor and effector organs).

To obtain a computational specification for temporal simulation of each state, a difference equation is used in the form of Eq. (4).

$$y_i(t + \Delta t) = y_i(t) + \gamma_i \left[g\left(\sigma, \tau, \sum_{j \in s(i)} \omega_{ji}y_j\right) - y_i(t) \right] \Delta t \quad (4)$$

By having different values for each parameter (i.e. for weight values ω_i, time step size Δt, slow and fast speed factors γ, steepness σ_i, threshold τ_i), the agent can facilitate a wide variety of behaviours. Each LP in Table 3 is represented in a computational form (by the JAVA language) and the dynamics of the system is achieved through evaluating the causality effects through a set of difference equations as in Eq. (4). For each discrete time step Δt the

Table 3 Specification of local properties in the hybrid language LEADSTO and in differential equation format

LP1	$EA(a_i, V_1)$ & $WS(s_k, V_2) \Rightarrow WS(s_k, V_2 + \gamma[f(\omega_1 V_1) - V_2]\,\Delta t)$
	$\dfrac{dWS(s_k)(t)}{dt} = \gamma[f(\omega_1.EA(a_i)(t)) - WS(s_k)(t)]$
LP2	$EO(Y_{(i,k)}, V_1)$ & $WS(c_k, V_2) \Rightarrow WS(c_k, V_2 + \gamma[f(\omega_2 V_1) - V_2]\,\Delta t)$
	$\dfrac{dWS(c_k)(t)}{dt} = \gamma\left[f\left(\omega_2.EO(Y_{(i,k)})(t)\right) - WS(c_k)(t)\right]$
LP3	$EA(a_i, V_1)$ & $EO(Y_{(i,k)}, V_2)$ & $WS(b_i, V_3) \Rightarrow WS(b_i, V_3 + \gamma[f(\omega_3 V_1, \omega_4 V_2) - V_3]\,\Delta t)$
	$\dfrac{dWS(b_i)(t)}{dt} = \gamma\left[f\left(\omega_3.EA(a_i)(t), \omega_4.EO(Y_{(i,k)})(t)\right) - WS(b_i)(t)\right]$
LP4	$WS(s_k, V_1)$ & $SS(s_k, V_2) \Rightarrow SS(s_k, V_2 + \gamma[f(\omega_5 V_1) - V_2]\,\Delta t)$
	$\dfrac{dSS(s_k)(t)}{dt} = \gamma[f(\omega_5.WS(s_k)(t)) - SS(s_k)(t)]$
LP5	$WS(c_k, V_1)$ & $SS(c_k, V_2) \Rightarrow SS(c_k, V_2 + \gamma[f(\omega_6 V_1) - V_2]\,\Delta t)$
	$\dfrac{dSS(c_k)(t)}{dt} = \gamma[f(\omega_6.WS(c_k)(t)) - SS(c_k)(t)]$
LP6	$WS(b_i, V_1)$ & $SS(b_i, V_2) \Rightarrow SS(b_i, V_2 + \gamma[f(\omega_7 V_1) - V_2]\,\Delta t)$
	$\dfrac{dSS(b_i)(t)}{dt} = \gamma[f(\omega_7.WS(b_i)(t)) - SS(b_i)(t)]$
LP7	$SS(s_k, V_1)$ & $SR(s_k, V_2) \Rightarrow SR(s_k, V_2 + \gamma[f(\omega_8 V_1) - V_2]\,\Delta t)$
	$\dfrac{dSR(s_k)(t)}{dt} = \gamma[f(\omega_8.SS(s_k)(t)) - SR(s_k)(t)]$
LP8	$SS(c_k, V_1)$ & $SR(c_k, V_2) \Rightarrow SR(c_k, V_2 + \gamma[f(\omega_9 V_1) - V_2]\,\Delta t)$
	$\dfrac{dSR(c_k)(t)}{dt} = \gamma[f(\omega_9.SS(c_k)(t)) - SR(c_k)(t)]$
LP9	$PA(a_i, V_1)$ & $SS(b_i, V_2)$ & $PO(Y_{(i,k)}, V_3)$ & $EO(Y_{(i,k)}, V_4)$ & $SR(b_i, V_5) \Rightarrow SR(b_i, V_5 + \gamma[f(\omega_{10}V_1, \omega_{11}V_2, \omega_{12}V_3, \omega_{13}V_4) - V_5]\,\Delta t)$
	$\dfrac{dSR(b_i)(t)}{dt} = \gamma\left[f\left(\omega_{10}.PA(a_i)(t), \omega_{11}.SS(b_i)(t), \omega_{12}.PO(Y_{(i,k)})(t), \omega_{13}.EO(Y_{(i,k)})(t)\right) - SR(b_i)(t)\right]$
LP10	$F(b_i, V_1)$ & $SR(s_k, V_2)$ & $PAwr(Y_{(i,k)}, V_3)$ & $PA(a_i, V_4)$ & $PA(a_i, V_5) \Rightarrow PA(a_i, V_5 + \gamma[f(\omega_{14}V_1, \omega_{15}V_2, \omega_{16}V_3, \omega_{17}V_4) - V_5]\,\Delta t)$
	$\dfrac{dPA(a_i)(t)}{dt} = \gamma\left[f\left(\omega_{14}.F(b_i)(t), \omega_{15}.SR(s_k)(t), \omega_{16}.PAwr(Y_{(i,k)})(t), \omega_{17}.PA(a_j)(t)\right) - PA(a_i)(t)\right]$
LP11	$SR(b_i, V_1)$ & $F(b_i, V_2) \Rightarrow F(b_i, V_2 + \gamma[f(\omega_{18}V_1) - V_2]\,\Delta t)$
	$\dfrac{dF(b_i)(t)}{dt} = \gamma[f(\omega_{18}.SR(b_i)(t)) - F(b_i)(t)]$
LP12	$SR(s_k, V_1)$ & $SR(c_k, V_2)$ & $PA(a_i, V_3)$ & $F(b_i, V_4)$ & $RO(Y_{(i,k)}, V_5)$ & $PO(Y_{(i,k)}, V_6) \Rightarrow PO(Y_{(i,k)}, V_6 + \gamma[f(\omega_{19}V_1, \omega_{20}V_2, \omega_{21}V_3, \omega_{22}V_4, \omega_{23}V_5) - V_6]\,\Delta t)$
	$\dfrac{dPO(Y_{(i,k)})(t)}{dt} = \gamma\left[f\left(\omega_{19}.SR(s_k)(t), \omega_{20}.SR(c_k)(t), \omega_{21}.PA(a_i)(t), \omega_{22}.F(b_i)(t), \omega_{23}.RO(Y_{(i,k)})(t)\right) - PO(Y_{(i,k)})(t)\right]$
LP13	$PO(Y_{(i,k)}, V_1)$ & $F(b_i, V_2)$ & $RAwr(Y_{(i,k)}, V_3)$ & $PAwr(Y_{(i,k)}, V_4) \Rightarrow PAwr(Y_{(i,k)}, V_4 + \gamma[f(\omega_{24}V_1, \omega_{25}V_2, \omega_{26}V_3) - V_4]\,\Delta t)$
	$\dfrac{dPAwr(Y_{(i,k)})(t)}{dt} = \gamma\left[f\left(\omega_{24}.PO(Y_{(i,k)})(t), \omega_{25}.F(b_i)(t), \omega_{26}.RAwr(Y_{(i,k)})(t)\right) - PAwr(Y_{(i,k)})(t)\right]$
LP14	$PA(a_i, V_1)$ & $PO(Y_{(i,k)}, V_2)$ & $PAwr(Y_{(i,k)}, V_3)$ & $EA(a_i, V_4) \Rightarrow EA(a_i, V_4 + \gamma[f(\omega_{27}V_1, \omega_{28}V_2, \omega_{29}V_3) - V_4]\,\Delta t)$
	$\dfrac{dEA(a_i)(t)}{dt} = \gamma\left[f\left(\omega_{27}.PA(a_i)(t), \omega_{28}.PO(Y_{(i,k)})(t), \omega_{29}.PAwr(Y_{(i,k)})(t)\right) - EA(a_i)(t)\right]$
LP15	$EA(a_i, V_1)$ & $F(b_i, V_2)$ & $PO(Y_{(i,k)}, V_3)$ & $SR(c_k, V_4)$ & $RO(Y_{(i,k)}, V_5) \Rightarrow RO(Y_{(i,k)}, V_5 + \gamma[f(\omega_{30}V_1, \omega_{31}V_2, \omega_{32}V_3, \omega_{33}V_4) - V_5]\,\Delta t)$
	$\dfrac{dRO(a_i\,b_i\,c_k\,s_k)(t)}{dt} = \gamma\left[f\left(\omega_{30}.EA(a_i)(t), \omega_{31}.F(b_i)(t), \omega_{32}.PO(Y_{(i,k)})(t), \omega_{33}.SR(c_k)(t)\right) - RO(Y_{(i,k)})(t)\right]$
LP16	$EA(a_i, V_1)$ & $RO(Y_{(i,k)}, V_2)$ & $F(b_i, V_3)$ & $PAwr(Y_{(i,k)}, V_4)$ & $RAwr(Y_{(i,k)}, V_5) \Rightarrow RAwr(Y_{(i,k)}, V_5 + \gamma[f(\omega_{34}V_1, \omega_{35}V_2, \omega_{36}V_3, \omega_{37}V_4) - V_5]\,\Delta t)$
	$\dfrac{dRAwr(Y_{(i,k)})(t)}{dt} = \gamma\left[f\left(\omega_{34}.EA(a_i)(t), \omega_{35}.RO(Y_{(i,k)})(t), \omega_{36}.F(b_i)(t), \omega_{37}.PAwr(Y_{(i,k)})(t)\right) - RAwr(Y_{(i,k)})(t)\right]$
LP17	$RO(Y_{(i,k)}, V_1)$ & $RAwr(Y_{(i,k)}, V_2)$ & $EO(Y_{(i,k)}, V_3) \Rightarrow EO(Y_{(i,k)}, V_3 + \gamma[f(\omega_{38}V_1, \omega_{39}V_2) - V_3]\,\Delta t)$
	$\dfrac{dEO(Y_{(i,k)})(t)}{dt} = \gamma\left[f\left(\omega_{38}.RO(Y_{(i,k)})(t), \omega_{39}.RAwr(Y_{(i,k)})(t)\right) - EO(Y_{(i,k)})(t)\right]$

Here $Y_{(i,k)}$ means (a_i, b_i, c_k, s_k) and for each LP, the first representation is in the LEADSTO and that will be followed by the differential equation format

behaviour of each state is calculated the emergence of the behaviour is traced with a identified parameter value set. From a mathematical point of view, the dynamics of the model is (numerically) solving the differential equations of LP1 to LP17 by assuming that at time $t = 0$, WS(s) and WS(c) holds value 1 as state level.

Some of these behaviours are presented as simulations in Sect. 4.

4 Simulation results

In this section, simulation experiments for a number of example scenarios are discussed, which all involve the occurrence of a preparation state for an action a_i, triggered by some stimulus s_k and context c_k. These scenarios relate to phenomena in the literature, as discussed in Sects. 1 and 2. They have been generated based on the specification presented in Sect. 3. Eight scenarios have been simulated to highlight the different possible behaviours of the model and among those three scenarios are new, whereas the other five were selected from the previous work in [22] but those behaviours have been significantly improved in the newer versions presented here. Furthermore, for the scope of this paper only c is 'self' situations are selected (for some examples where c is 'other' see [22, 23]). The following is a summary of the different simulated scenarios:

(1) The first scenario simulated describes a situation where the prepared action has satisfactory predicted effects and therefore is executed; in this case both prior and retrospective awareness states occur. This scenario will be considered as the base case for the interplay between conscious and unconscious processes.

(2) The second scenario simulated describes a situation where the prepared action has satisfactory predicted effects and therefore is executed but the awareness is absent (in other words merely an unconsciousness action). The strength of the action execution is lower as compared to the first scenario.

(3) The third scenario simulated describes a situation where the prepared action lacks satisfactory predicted effects, and is therefore not executed: a no–go decision, or vetoing in unconscious form. Furthermore, the awareness state is almost absent due to the almost absent feeling.

(4) The fourth scenario simulated describes a situation where a poor action prediction capability is modelled: the action effect is falsely predicted as satisfactory. This leads to a prior ownership state, which is sufficient to actually execute the prepared action. In this case, a low retrospective ownership

state and almost absent retrospective awareness state will occur, as the sensory representation of the effect stays low. This simulation is used to explain the basic cognition and behaviour of a schizophrenic patient.

(5) The fifth scenario simulated describes a situation where two prepared actions exist for two input stimuli (s_1, s_2, c_1 and c_2) but one is relatively less positive compared to the other on predicted effects (difference is 0.2 in terms of the weight ω_{10}). The one which is less positive (2nd option) gets diluted over time in terms of PA(a_2), SR(b_2), F(b_2), PO(a_2, b_2, c_2, s_2) and PAwr(a_2, b_2, c_2, s_2) while the other prepared action gets executed and develops the retrospective awareness too.

(6) The sixth scenario simulated describes a situation exactly as in the fifth scenario but in this case once the action with the strongest predictive effect [i.e. SR(b_1)] is executed it does not suppress the inputs s_2 and c_2. Therefore, once EA(a_1) is executed because of the existence of s_2 and c_2 the agent is preparing for EA(a_2) and successfully this will be performed. This confirms the agent capacity of cognitive control combining both conscious and unconscious processes.

(7) The seventh scenario simulated describes a situation where the agent is prepared for an action by expecting a particular effect b_1, though it is actual effect after execution is different: b_2 (mismatch between the predicted and actual). As claimed by Haggard, this scenario contributes to the idea that awareness of an action is a dynamic combination of both predictive and inferential sense-making. Having a strong predicted effect agent develops a strong prior awareness, but not sensing the same effect it leads to a poor retrospective awareness of that predicted effect. This phenomenon is important for the agent's learning process through error correction.

(8) The eighth scenario simulated describes a situation that can be considered as an early stage of a cognitive impairment for depressive symptoms. In this case, the agent is preparing for two action options where one is having a positive feeling while the other is with a negative feeling [i.e. F(b_1) and F(b_2)]. According to the biased nature on negative feelings, the agent consciously affects selection of the negative one, though both options are identical from an action selection perspective (i.e. by having exactly the same values for all the parameters for each option except for ω_{29}). Due to this conscious biased influence agent will execute the action with a negative feeling and it is considered to be that repeatedly performing these type of thoughts/tasks it

will develop a negative mood that leads to a depression situation (in a long run).

Although these simulations will be presented in more abstract manner without relating them to real examples all the time, in general from a real-world perspective, the following example explanation of a scenario can be kept in mind:

- Stimulus s is that you need to withdraw some money.
- Context c is that you are in a bank and in front of an ATM.
- Action a is clicking on specific buttons to withdraw € 55.
- Effect b of action a is that you get € 55 cash.
  This simple happy example scenario may have many variations such as:
- the person may doubt whether to directly withdraw € 55 or to first check the current balance in the account
- once the person has entered € 55 to withdraw, the system may inform the person that it does not have € 5 notes and ask to change the amount
- may be surprisingly ATM will return the card but without the money

Furthermore, when considering complex systems like Air Traffic Management there will be many examples that this model can be utilised (see [58] for more examples).

4.1 Selecting values for parameters of the model

Selecting parameter values for a dynamic cognitive model which consists with set of differential equations is a non-trivial research challenge. This will be even more difficult when there are no detailed numerical empirical data to use in this process, but only some characteristics of behaviours available in more fuzzy form (cf. [59]). Furthermore, another problem with computational cognitive models is that for different types of persons with different behaviours it may seem necessary to find a unique person-dependent set of parameters from scratch. As an alternative, if it is possible to identify a particular parameter value set which is able to demonstrate a variety of situations using only very minimum number of variations this would make the issue easier to handle. Therefore, in this work the focus has been on finding such a generic parameter value set that provides more confidence from the model validation perspective and its practical usage in future complex applications. The current model consists of many parameters: 39 weight values (for one option: $k = i = 1$): ω_i, a time step size: Δt, slow and fast speed factors: γ, 17 steepness: σ_i, and 17 threshold: τ_i. Table 4 presents connection weight values and Table 5 presents threshold (τ) and steepness (σ) values used in configurations of simulations on this cognitive model. From

these it is clear that the weight value set is generic and just changing very few weights (in most of the cases either one or two) have obtained the different expected behaviours.

The main challenge in this approach is that there is no real detailed data value set that can be compared to the output of the agent model to estimate parameters. Only certain features of the behaviour of each cognitive state are known for different scenarios, based on neurological and behavioural evidence from the literature (for example prior awareness should occur before the action execution and after prior ownership, there should be a dip in the sensory representation in-between predictive representation and inferential representation, et cetera). To identify the parameter values a systematic approach is used. For this approach, it is a necessary condition to select multiple scenarios (minimum is three but having more will improve the quality of the results) which are interrelated from a functional point of view. For example, the first scenario is considered to be as a reference scenario and the second scenario is different from that just by achieving that awareness states are not developing (i.e. $\omega_{24} = \omega_{25} = \omega_{34} = \omega_{35} = \omega_{36} = 0$). Furthermore, the third scenario handles a case in which a prepared action lacks satisfactory predictive effects (i.e. $\omega_{24} = 0.2$), and the fourth scenario addresses poor action prediction capability (i.e. $\omega_{18} = 0.2$); and so on (see Table 4). This interrelation among scenarios is very important for a minimum number of parameter changes enabling to identify a generic parameter value set for the model.

In this parameter estimation approach the idea is as follows:

- First scenario is addressed and parameter values are calibrated to simulate its behaviour as identified through the literature.
- Then using the obtained parameter value set, by changing just a few (scenario-related) weight values it is checked whether the model with these parameter settings is able to generate the behaviour for the second scenario.
- If this provides a simulation with a pattern as expected (without changes to the previously obtained parameter values, except for the changes particular to the current scenario) then it provides a good confidence on the current identified parameter value set,
- But if not, then it is necessary to change the parameter values of the first simulation (based on the sensitivity of certain parameters on the required final output) until the behaviours for both simulations are satisfactory.
- This approach is incrementally extended to each scenario until a generic parameter value set for all the scenarios has been identified. For any new scenario if any changes to the previously obtained parameter values are required, then all previously addressed scenarios are readdressed.

Table 4 Connection weight values used for cognitive agent model

Weights	Scen. 1	Scen. 2	Scen. 3	Scen. 4	Scenario 5		Scenario 6		Scenario 7		Scenario 8	
					Stim. 1	Stim. 2	Stim. 1	Stim. 2	Effect1	Effect2	Stim. 1	Stim. 2
ω_1	0.5				*	*				–	*	*
ω_2	0.8				*	*			*	–	*	*
ω_3	0.7								0.0	0.5		
ω_4	0.7								0.0		*	*
$\omega_5 - \omega_6$	1.0									–		
ω_7	0.8									0.5		
$\omega_8 - \omega_9$	1.0									–		
ω_{10}	0.8	0.2				0.6		0.6		0.0		
ω_{11}	0.8											
ω_{12}	0.5									0.0		
ω_{13}	0.9											
$\omega_{14} - \omega_{15}$	0.8									–		
ω_{16}	0.7									–		
ω_{17}	0.0				0.8	0.8	0.8	0.8		–	0.1	0.1
ω_{18}	0.8		0.2									
ω_{19}	0.3									–		
$\omega_{20} - \omega_{22}$	0.8									–		
ω_{23}	0.8									–		
ω_{24}	0.9	0.0								–		
ω_{25}	0.8	0.0								–		
ω_{26}	0.8									–		
$\omega_{27} - \omega_{28}$	0.8									–		
ω_{29}	0.5									–	0.1	0.8
ω_{30}	0.8								0.0			
ω_{31}	0.8											
ω_{32}	0.8									0.0		
ω_{33}	0.8											
ω_{34}	0.7	0.0							0.0			
ω_{35}	0.7	0.0										
ω_{36}	0.7	0.0										
ω_{37}	0.7									0.0		
$\omega_{38} - \omega_{39}$	0.8											

In here if a value of a particular weight is empty for a scenario that means it is equal to the value of that in the Scenario 1. Furthermore if a value is '–' then such a link was not existed for that scenario and furthermore '*' presents that the particular link suppresses both its mapping inputs and all the remaining

- In the first few iterations, it may challenging to identify a parameter value set, but over time the convergence is really fast and it will be experienced that the identified parameter value set is more generic and it is facilitating the necessary behaviours for simulations even without any changes to the obtained parameter value set.

In addition to the parameter values in Tables 4 and 5 for the step size (Δt), slow speed factor (γ), and fast speed factor (γ) parameter values 0.25, 0.6 and 0.7 were used respectively for all the scenarios.

4.2 Scenario 1: normal execution with ownership and awareness

The first scenario considered describes a situation where the context c is the agent itself, and a stimulus s occurs.

The action effect b of a is considered positive for the agent and the awareness of action formation and execution will occur, together with generated prior and retrospective ownership states. The following execution trace will be expected from the agent here:

- External stimulus s and context c will occur and trigger preparation of action a.
- Based on the preparation state for a, the sensory representation of the (positive) predicted effect b of a is generated.
- Based on this positive predicted effect and the other states, a prior ownership state for action a is generated.
- Prior ownership for action a is followed by the prior awareness; this is generated just before the action execution.

Table 5 Threshold (τ) and Steepness (σ) values used in configurations of simulations

	WS(s_k)	WS(c_k)	WS(b_i)	SS(s_k)	SS(c_k)	SS(b_i)	SR(s_k)	SR(c_k)	SR(b_i)
σ	1.00	1.00	6.00	2.00	2.00	3.50	2.00	2.00	3.00
τ	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01

	PA(a_i)	F(b_i)	PO($Y_{i,k}$)	PAwr($Y_{i,k}$)	EA(a_i)	RO($Y_{i,k}$)	RAwr($Y_{i,k}$)	EO($Y_{i,k}$)
σ	2.00	3.50	6.00	5.00	6.00	5.00	4.00	5.00
τ	0.50	0.01	1.00	0.60	1.00	1.90	1.60	0.80

– The prior ownership and prior awareness states for action *a* lead to actual execution of action *a*.
– The execution of *a* affects *b* in a positive manner and propagates to the sensory representation of *b* and the feeling of *b*.
– At the same time the sensory representation of *b* is suppressed due to the prior self ownership state.
– Based on the generated states, after the execution of action *a* the agent develops a retrospective ownership state.
– Retrospective ownership of sensed effect *b* of action *a* is followed by retrospective awareness of action *a* and its effect
– Finally, the agent communicates this ownership and awareness of it.

The simulation result of this scenario is shown in Fig. 2. In this figure, it is shown that (after sensing the stimulus) the agent triggers preparation of action *a* from time point 3 on (with a peak value of 0.74 around time point 55). Based on that the sensory representation of the predicted effect *b* of *a* is generated (through the as-if body loop with peak value 0.15 around time point 15 and through the body loop with peak value 0.59 after action execution around time point 50). This is followed by the feeling of *b* (through the as-if body loop with the peak value 0.19 and through the body loop with the peak value 0.62). These states contribute to the generation of a prior ownership state which starts to trigger at time point 5 and reaches a peak value of 0.77 around time point 57. After activating prior ownership, prior awareness is developing, mainly upon the formation process of effect prediction *b* of *a*, and its associated feeling. The prior awareness has started to pop up around time point 13 and has obtained peak value 0.71 around time point 45. As a result of the prior ownership and awareness states, the agent initiates the actual execution of action *a* which propagates its effects through the (external) body loop. This clearly shows that prior awareness is just before the action execution (cf. [4, 5, 8]) as the action execution process started at time point 15, and has its peak around time point 55, with maximum strength 0.92. Furthermore, it shows that via the body loop and the sensor state, the execution of action

a also affects the sensory representation of *b* and the feeling of *b*. Therefore, the sensory representation *b* of *a* behaves as expected, by adding the sensed actual effect to the predicted effect, and the same effect is propagated to the feeling of *b* too (cf. [20, 47, 48]). Due to the action execution, the agent develops a retrospective ownership state (starts at time point 26 with peak value 0.82 around time point 58), which is followed by a retrospective awareness state (starts at time point 29 with peak value 0.78 around time point 62). Finally, the agent communicates ownership and awareness of it for the performed action, based on the retrospective awareness and ownership states (with the maximum strength of 0.82 around time point 67). Note that when the stimulus is taken away (as explained in Sect. 3 this has been performed through an external suppressive mechanism through the orange colour arrows in Fig. 1), all activation levels will come down to 0; they will come up again when a new stimulus occurs. Note that the numerical information related to the time scale or the peak values has not been coupled with actual brain signals but is only used as a frame of reference.

4.3 Scenario 2: normal execution with ownership but without awareness

There are many situations in which human action formation occurs merely unconsciously, especially when related to habitual tasks [40]. As in the first scenario above also in this case, the agent will experience that the prepared action has satisfactory predicted effects. Nevertheless, the agent will not develop any awareness state of the experienced feeling. The following execution trace will be expected from the agent for this scenario.

– External stimulus *s* and context *c* will occur and trigger preparation of action *a*.
– Based on the preparation state for *a*, the sensory representation of a (positive) predicted effect *b* of *a* is generated.
– Based on this positive predicted effect and the other states a prior ownership state for action *a* is generated, but no prior awareness
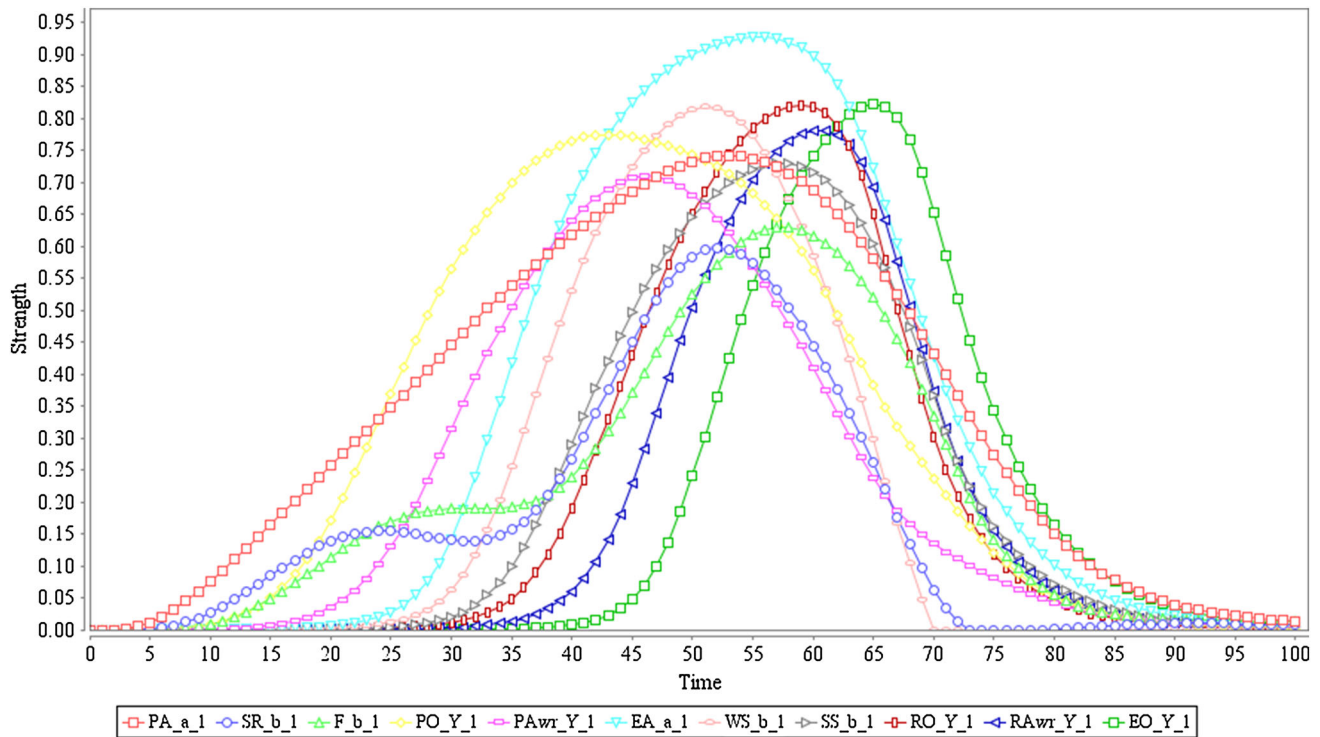
Fig. 2 Scenario 1: Executing an action with ownership and awareness. In here 'Y_1' represents 'a_1, b_1, c_1, s_1'

- The prior ownership state for action a leads to actual execution of action a.
- The execution of a affects (via sensing) the sensory representation of b and the feeling of b in a positive manner
- At the same time, the sensory representation of b is suppressed due to the prior ownership state.
- Based on the generated states, after the execution of action a, the agent develops a retrospective ownership state, but no retrospective awareness.
- The agent does not communicate this ownership.

The simulation result for this process is presented in Fig. 3. The agent starts to prepare for action a at time point 2, and for this preparation a peak value of 0.52 is obtained. Together with the action preparation, agent develops the sensory representation of predicted effect b of a (with peak value 0.15 based on the as-if body loop, and peak value 0.61 through the body loop) and the associated feeling of b (with peak value 0.68). Based on this predictive information, the agent develops prior ownership from time point 6 on and with peak value 0.7. More importantly, in this simulation prior awareness has not developed. The developed states lead to performing the action a which starts at the time point 14 and obtains peak value 0.43. The execution positively affects (via the sensor state) the sensory representation of b and the feeling of b (adding the sensed actual effect to the predicted effect). With this action execution effect, the agent develops retrospective

ownership with peak value 0.44. Finally, the agent does not communicate the ownership about the performed action (has a very low strength due to lack of awareness).

This cognitive behaviour trace is in line with the expectations, and mainly when comparing it to Scenario 1, it demonstrates the possible impact of the awareness states. For example, when a prior awareness state occurs, it facilitates an enhancing effect on action preparation, sensory representation and feeling and also provides much smoother effects (as seen in Fig. 2). Nevertheless, in this Scenario 2, these states have lower activation levels (also the action preparation state), that may be attributed to the absence of enhancing effects through prior awareness. Also this scenario highlights the strength of action execution with awareness: the action has executed with a relatively high peak value (0.92 vs. 0.43). This might be explained as an influence from prior awareness to action execution. Nevertheless, according to the literature, further research is required to conclude this; cf. [5]. From the computational perspective, this at least confirms the model's capability of action formation both with awareness and without awareness.

4.4 Scenario 3: prepared action lacks satisfactory predicted effects

Humans are not always responsive to all the environmental stimuli (even in unconscious form). Nevertheless, there should be an explanation from the perspective of internal
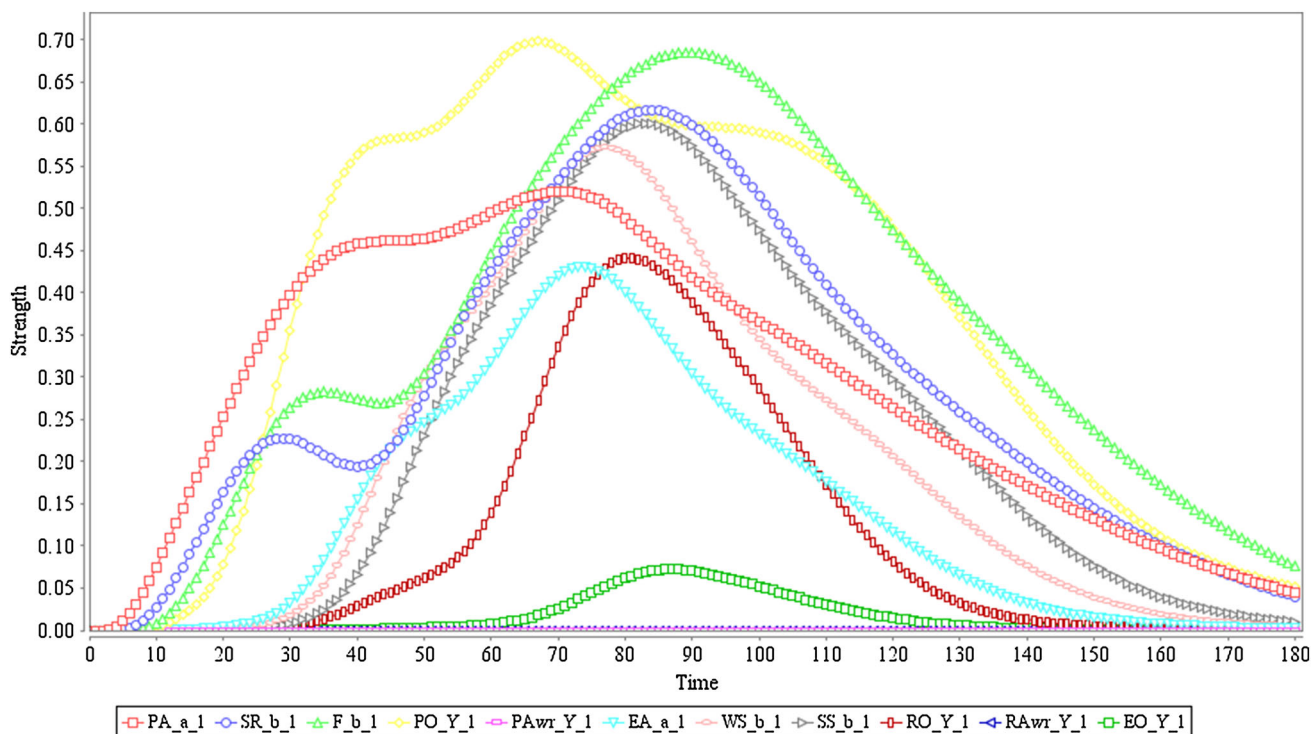
Fig. 3 Scenario 2: Executing an action with ownership and no awareness. In here 'Y_1' represents 'a_1, b_1, c_1, s_1'

processes when receiving a stimulus why that stimulus does not lead to an actual action execution. This simulation provides the behaviour for such situation and explains how the lack of (positive) predicted effects generated through the as-if body loop relates to this. The following execution trace will be expected for this scenario:

– External stimulus *s* and context *c* occur and trigger preparation of action *a*.
– Based on the preparation state for *a*, only a weak sensory representation of predicted effect *b* of *a* is generated.
– Poor predictive effects will be reflected through the feeling state with very low activation.
– Based on this poor predicted effect *b* and the other states, a low prior ownership state for action *a* is generated.
– Due to the poor prior ownership and low predictive feeling states, the agent does not develop adequate prior awareness.
– The low prior ownership state for *a* does not lead to actual execution of action *a*; the action *a* can be considered vetoed
– The agent develops no retrospective ownership state for *a* and no retrospective awareness.
– The agent does not communicate ownership or awareness.

The simulation of this scenario is shown in Fig. 4. The predicted effect is very low compared to the Scenario 1. This

clearly shows that the action *a* triggered by stimulus *s* (which has an effect *b*) is not positive for the agent (in other words it is more like neutral to the agent in terms of the feeling): it leads to not getting any feelings out of it. Nevertheless, the prediction capabilities are assumed correct in this case, so no high level of *b* is correctly predicted for *a*. As a result of this low prediction, the prior ownership state also stays at a very low level. Due to this, prior awareness is not developed (stays in a very low level), which would be needed to strengthen the action execution. Therefore execution of the action also stays very low (below 0.1) and due to that, there is no retrospective ownership state, nor communication of ownership. Having a single weight value change [ω_{10}: PA(a_i) to SR(b_i) from 0.8 to 0.2] to obtain this behaviour from the reference Scenario 1 shows the coherent nature of action formation and higher order coupling as a process. This shows, from a complex systems simulation perspective, how the same agent model by limited variations in assigned parameter values demonstrates qualitatively different results. Furthermore, this confirms that the model has adequately adapted Damasio's hypothesis: in the agent's decision-making process, it has to assess the incentive value of the choices through an internal simulation process.

4.5 Scenario 4: poor feelings of action prediction effects of a schizophrenic patient

In the previous scenario, it was presented how the lack of satisfactory predicted effects will lead to a No–Go decision
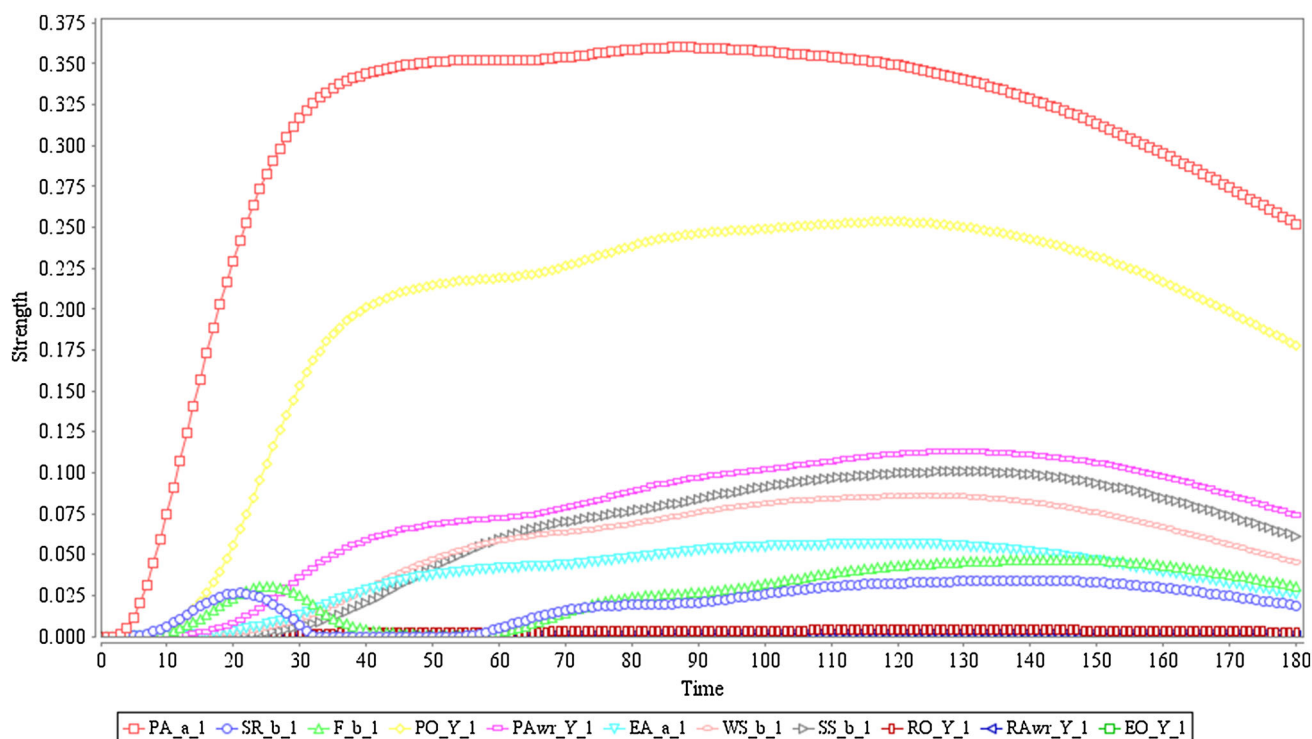
Fig. 4 Scenario 3: Prepared action lacks satisfactory predicted effects. In here 'Y_1' represents 'a_1, b_1, c_1, s_1'

on action execution. In the current simulation scenario, the focus is on poor feelings of predicted effects (the satisfaction level of the predicted effect is high but it does not adequately feel to the agent). In this situation the action effect b for action a, in principle is positive for the agent, like in the first scenario. Nevertheless, the agent will not properly feel the effects of prediction. This is what assumed to happen in (at least some) patients with Schizophrenia [38, 47, 48]. Much evidence exists that relates this to poor emotional aspects in expression, experience and perception (mainly due to abnormalities in the workings of Amygdala) [60, 61]. Schizophrenic patients often have the impression that their own actions are being created, not by themselves, but by someone from the outside [38]. The following execution trace will be expected for such a phenomenon:

- External stimulus s and context c occur and trigger preparation of action a.
- Based on the preparation state for a the sensory representation of predicted effect b of a is generated.
- Lack of feeling of predicted effects will be experienced.
- Based on this predicted effects and its poor feeling, a relatively low level of a prior ownership state for action a is generated.
- Based on this low level of prior ownership and poor feelings, a relatively low level of prior awareness will be developed.

- This prior ownership and awareness levels for action a are still sufficient to lead to actual execution of action a.
- The execution of a affects b in a positive manner and (via sensing) the sensory representation of b but still the felt feelings are weaker.
- Due to poor feelings, agent will not develop adequate level of retrospective ownership and retrospective awareness.
- The agent does not communicate ownership or awareness for action a.

The simulation of this scenario is shown in Fig. 5. In this case, the agent has not fully felt the predicted effects of action a. After sensing, the stimulus agent has triggered preparation of action a at time point 3 (with a peak value of 0.52). Based on that the sensory representation of predicted effect b of a is generated (through the as-if body loop with peak value 0.25 and through the body loop with peak value 0.59) and followed by the feeling of b (through the as-if body loop with peak value 0.08 and through the body loop with peak value 0.25). This clearly shows that the predicted effect has not been properly felt by the agent due to very low values for feeling state $F(b_i)$. Next these states contribute to generate a prior ownership which starts to trigger at the time point 6 with peak value 0.52. Together with the prior ownership, the agent has experienced prior awareness with peak value 0.47 (this strength is relatively less
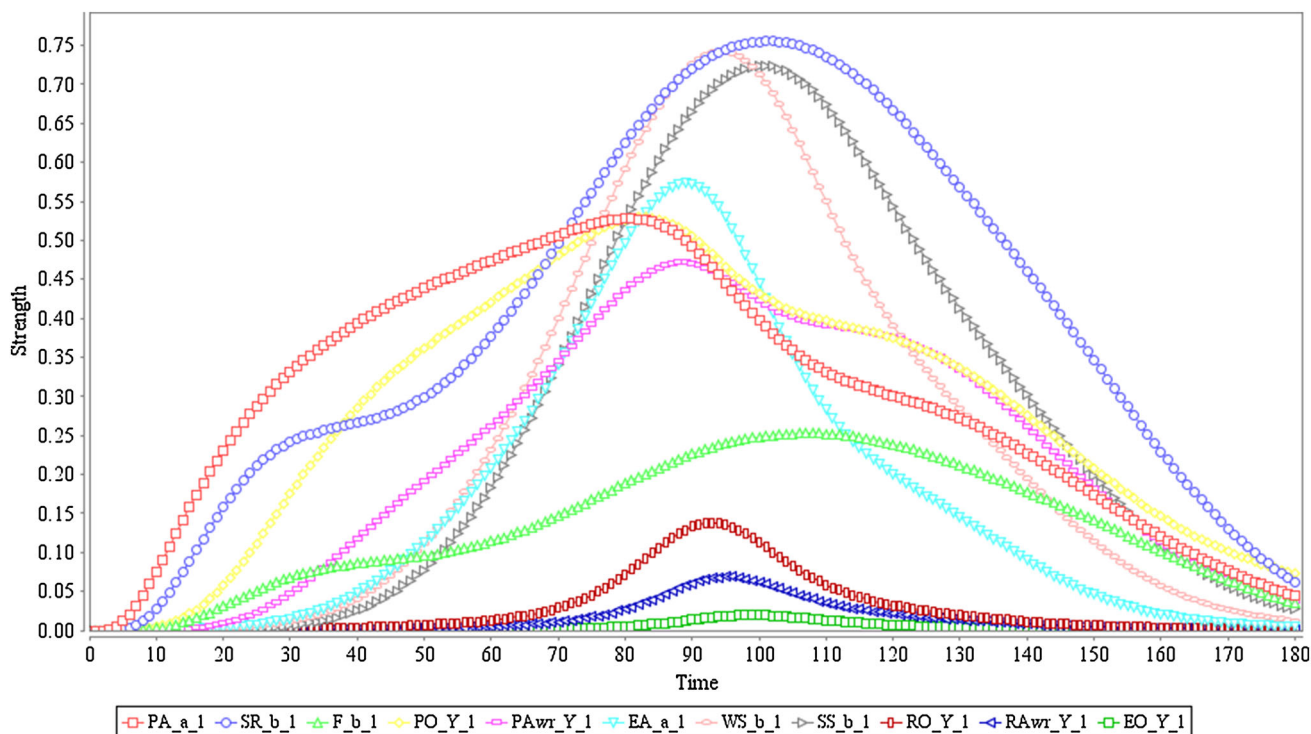
Fig. 5 Scenario 4: Poor feelings of action prediction effects of a schizophrenic patient. In here 'Y_1' represents 'a_1, b_1, c_1, s_1'

compared to the same in the first scenario: 0.71). The prior ownership and awareness levels are much better in this case compared to the situation in the third scenario. Therefore, in contrast to the third scenario, these levels turn out high enough for the execution of the action. The maximum strength of the actual execution of action a is 0.57 and this execution has positive effects which are sensed. Therefore, the sensory representation b of a behaves as expected after adding the sensed actual effect to the predicted effect. Nevertheless, the agent has again not properly felt the effects: it has not developed a sufficient strength for the feeling state. Due to these poor perceived effects of feeling, the agent has not developed an adequate level of retrospective ownership and no retrospective awareness. This behaviour can be interpreted as having some strength for the prior ownership and awareness for the action but no retrospective values for it. The agent may reach an internal conflict situation where the action seems not being created by itself, but by someone else. From the parameter values' perspective, this result has been achieved only with a single change from the first scenario on ω_{18} [from SR(b_i) to F(b_i)] from 0.8 to 0.2. Nevertheless, the cognitive impairment behind a schizophrenic patient is more complicated than modelled here; for example, the impact of some other states (perception, attention, emotions, etc.) has to be considered as well. Therefore, it is required to extend the current model to provide more

realistic cognitive behaviour for Schizophrenia. Nevertheless, the current behaviour already demonstrates some of the the basics.

4.6 Scenario 5: cognitive controlling when multiple action options compete

The fifth simulation scenario explains a situation when two action options exist (for the simulation two input stimuli were considered for this) and both have the potential to execute an action. The competition among the two options and the cognitive control through unconscious and conscious processes on this action selection is captured, resulting in execution of only one of the actions. For this behaviour two independent input tuples (s_k, c_k) were used, occurring in parallel (nevertheless it is possible to use a single input that triggers two action options, but due to the requirements in the next scenario the mentioned approach was used). For each option, the same configurations were used as in the first scenario, except for few weights (see Table 4). In this scenario, for the connection strength related with option 2 (i.e. $k = i = 2$) from the action preparation a_2 to its predicted effect b_2 a moderately low value has been selected: $\omega_{10} = 0.6$. Values for the other parameters were again the same as in Scenario 1 (cf. Table 4). More importantly, if in this situation at a given time only one input tuple occurs (either s_1, c_1 or s_2, c_2) then
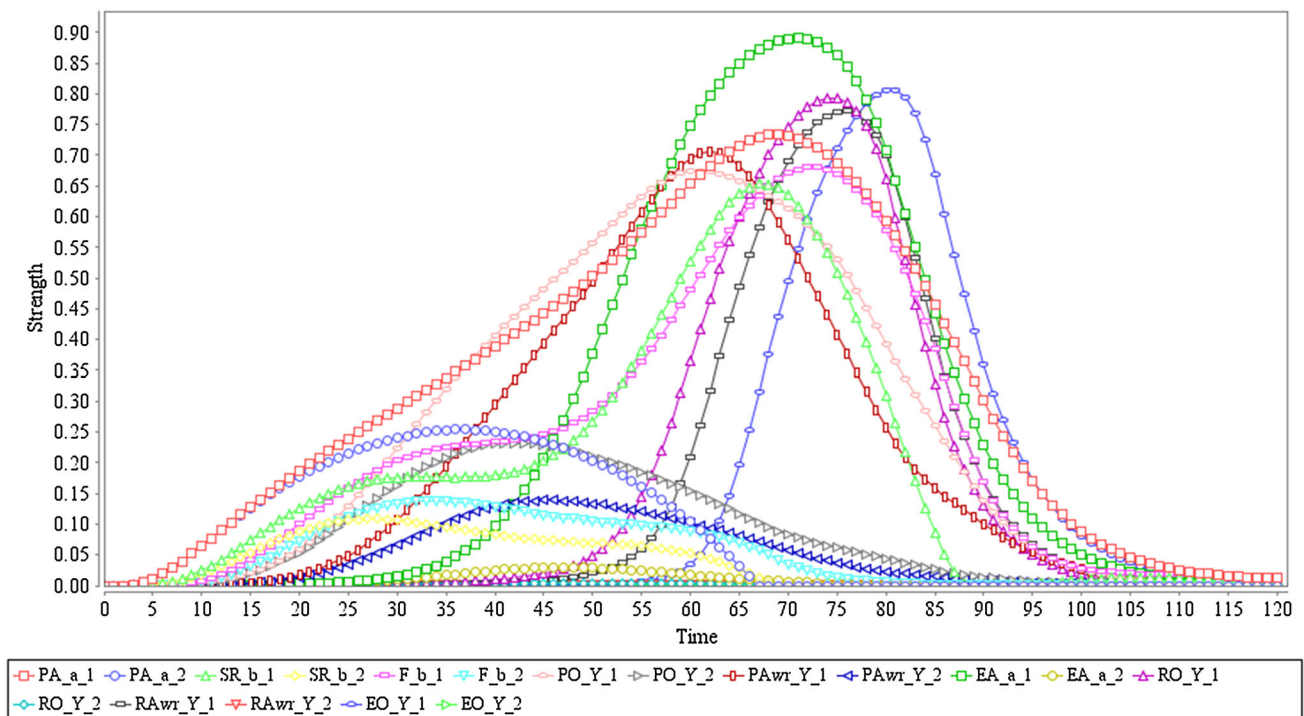
Fig. 6 Scenario 5: Cognitive controlling when multiple action options compete. In here 'Y_1' represents 'a_1, b_1, c_1, s_1' whereas 'Y_2' represents 'a_2, b_2, c_2, s_2'

each action will execute according to the same behaviour as in the first scenario. The simulation of this scenario is shown in Fig. 6. The following execution trace is expected from the agent in this case:

- External stimuli s_1, s_2 and contexts c_1, c_2 will occur and trigger preparation of actions a_1 and a_2 separately.
- Based on the preparation state for a_1 and a_2, the sensory representations of predicted effect b_1 of a_1 and b_2 of a_2 are also generated.
- Nevertheless, the agent will show strong effects on option a_1 and the rate of activation for preparation of a_2 will quickly slow down and disappear subsequently, due to suppression by the preparation for the other option.
- Based on this positive predicted effect and the other states for a_1, a prior ownership state for action a_1 is generated.
- Prior ownership for action a_1 is followed by the prior awareness, generated just before the action execution.
- The agent will develop neither prior ownership nor awareness for option a_2.
- Prior ownership and prior awareness states for action a_1 lead to actual execution of action a_1.
- The execution of a_1 affects b_1 in a positive manner and, via sensing, the sensory representation of b_1 and the feeling of b_1.

- At the same time the sensory representation of b_1 is suppressed due to the prior self ownership state of it.
- Based on the generated states, after the execution of action a_1 the agent develops a retrospective ownership state for action a_1.
- Retrospective ownership of action a_1 is followed by retrospective awareness of action a_1.
- Finally, the agent communicates this ownership and awareness related with option a_1.

The behaviour captured in the Fig. 6 is in line with the expected trace. Both action preparations [i.e. PA(a_1) and PA(a_2)] are activated at time point 2 and, more importantly, with the same rate of activation strength until time point 15 (this highlights the non-biased effects in early action formation). Nevertheless, after time point 15, it is clear that option a_1 has maintained somewhat the same rate of activation for action preparation, while for option a_2 the preparation activation strength has started to decrease, due to suppression by the preparation for option a_1, via the suppressive link with strength ω_{17}. In parallel to action preparation, the sensory representations and feelings for each option are also activated, but due to the assigned slightly lower value for weight link ω_{10} on option 2 the sensory representations and feelings for predicted effect of option 2 are not maintained, in addition to the effect through the suppressive link ω_{17} On the preparation state

for a_2. Each option for action preparation independently suppresses its complements', proportional to the current strength of each preparation (as in [51] for lateral inhibition). As ω_{10} for option 2 is slightly weaker, this contributes to the relatively a low value of PA(a_2) in comparison with PA(a_1). Therefore, through these unconscious mechanisms, the activation level of preparation state PA(a_2) becomes lower. Due to this bias in the action formation process, none of the other states related to the option 2 are activated. In contrast, for option 1 all the remaining states are activated in the same order as in the first scenario. More importantly, not having strong dips for sensory representation and feeling states as in the second scenario, this further highlights the conscious influence for action formation when compared to the aspects highlighted for the second scenario. The agent has developed both prior and retrospective awareness states with acceptable strength and finally has communicated the ownership and awareness specific to option 1.

As presented in the third section, this model includes some suppressive external links for the purpose of the scenario; for this scenario, once the action related to option 1 was executed it has suppressed all of the inputs, even those related to option 2 (i.e. s_2, c_2) and therefore all the states values become zero at the end of the simulation.

4.7 Scenario 6: cognitive control when multiple action options compete: an extension of the fifth scenario

In the fifth scenario once the action a_1 related to option 1 was executed, it has suppressed (or stopped) the input stimuli related with both the options (i.e. s_1, c_1, s_2 and c_2). In the current scenario, the same identical setup was used but execution of action a_1 does not stop the input stimuli related with the second option. As the action option 2 has been suppressed by action option 1, once the effects of action option 1 have been realised, and due to that its triggers have disappeared, execution of action option 2 can get a second chance in the action formation process. The following execution trace will be expected from the agent in this case:

- External stimuli s_1, s_2 and contexts c_1, c_2 will occur in parallel and trigger preparation of actions a_1 and a_2.
- Based on the preparation states for a_1 and a_2, the sensory representations of predicted effect b_1 of a_1 and b_2 of a_2 are generated.
- Nevertheless, the agent will show stronger effects on option a_1 and the rate of activation for a_2 will quickly slow down and disappear subsequently.
- Based on the positive predicted effect and the other states for a_1 a prior ownership state for action a_1 is generated.

- Prior ownership for action a_1 is followed by prior awareness; this is generated just before the action execution.
- The agent develops neither prior ownership nor awareness for option a_2.
- Prior ownership and prior awareness states for action a_1 lead to actual execution of action a_1.
- The execution of a_1 affects b_1 in a positive manner and, via sensing the sensory representation of b_1 and the feeling of b_1.
- At the same time, the sensory representation of b_1 is suppressed due to the prior self ownership state of it.
- Based on the generated states, after the execution of action a_1, the agent develops a retrospective ownership state for action a_1.
- Retrospective ownership for action a_1 is followed by retrospective awareness of action a_1.
- Finally the agent communicates this ownership and awareness related with option a_1.
- With the execution of action a_1, the agent suppresses the inputs s_1 and c_1.
- Due to the absence of inputs s_1 and c_1, preparation for action a_1 is not triggered anymore.
- The still existing inputs s_2 and c_2 still trigger preparation of action a_2. This is not suppressed by preparation of action a_1, since this is not activated anymore.
- Based on the preparation state for a_2 the sensory representation of predicted effect b_2 of a_2 is generated.
- Based on this positive predicted effect and the other states a prior ownership state for action a_2 is generated.
- Prior ownership for action a_2 is followed by prior awareness for action a_2, which is generated just before the execution of action a_2.
- Prior ownership and prior awareness states for action a_2 lead to actual execution of action a_2.
- The execution of a_2 affects b_2 in a positive manner and, via sensing the sensory representation of b_2 and the feeling of b_2.
- At the same time, the sensory representation of b_2 is suppressed due to the prior self ownership state.
- Based on the generated states, after the execution of action a_2 the agent develops a retrospective ownership state for action a_2.
- Retrospective ownership of action a_2 is followed by retrospective awareness of action a_2.
- Finally the agent communicates this new ownership and awareness too.

The simulation of this scenario is shown in Fig. 7. In this figure, from time point 0 to (roughly) 80 the behaviour is exactly the same as in the fifth scenario, but around time point 80 due to the suppressive effect on inputs s_1 and c_1 the agent shows the effects of losing the suppression of action
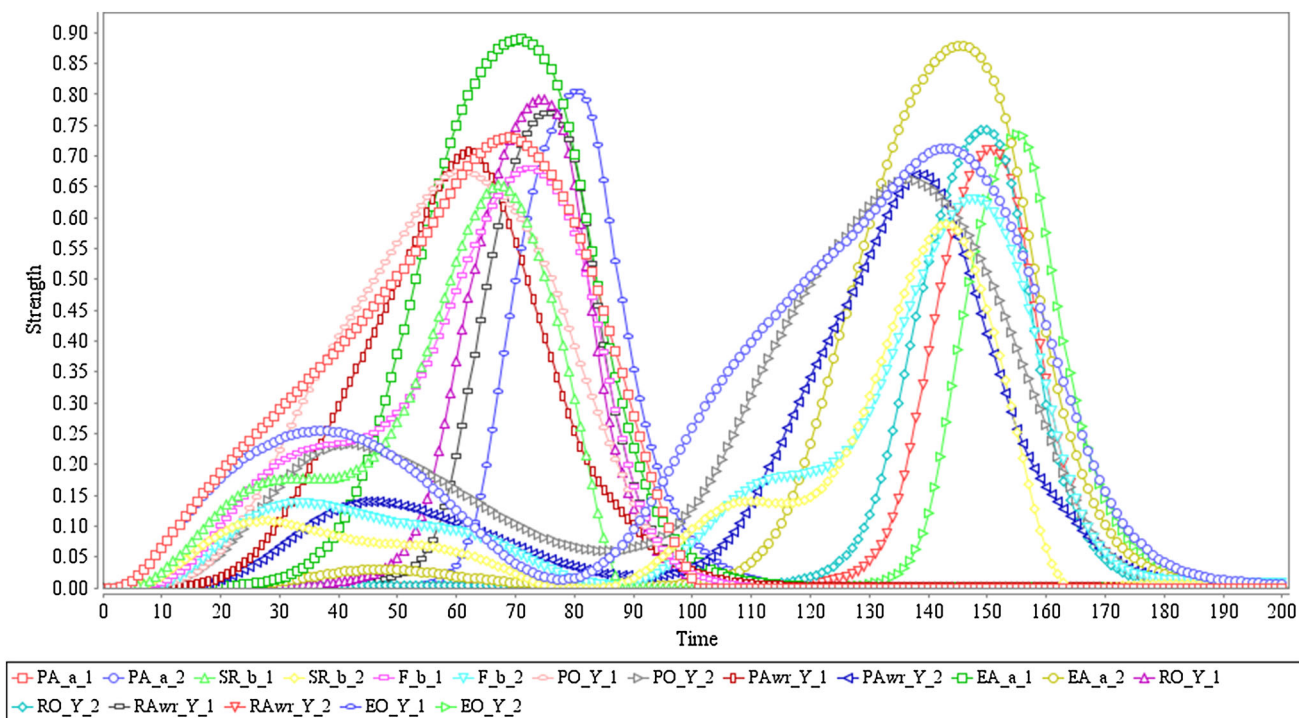
Fig. 7 Scenario 6: Effects of cognitive controlling when multiple action options compete but after the first action execution. In here 'Y_1' represents 'a_1, b_1, c_1, s_1'

option a_2. Therefore, the agent has again strengthens the preparation of action a_2 and subsequently the sensory representation of predicted effect b_2 of a_2, and the feeling of b_2. Prior ownership co-occurs with the above states (mainly due to its pre obtained activation strength). Followed by the prior ownership state, the prior awareness state develops as expected. As a result of the prior ownership and awareness states, the agent initiates the actual execution of action a_2, which propagates its effects through the external body loop. This shows a possibility to get action a_2 executed, in contrast to Scenario 6 above where due to the action competition it is suppressed. The peak value obtained for the action execution is 0.88 (the same for option a_1 is 0.89); this clearly shows that both options have the same power in getting executed. The execution of action a_2 (via the body loop) has further effects too. Due to the action execution, the agent develops the retrospective ownership state for action a_2, which is followed by a retrospective awareness state. Finally, the agent communicates ownership and awareness about the just performed action. This simulation shows the ability of action selection through competition mainly from an unconscious perspective. Having a slightly different weight value for ω_{10} [from PAwr(a_i, b_i, c_k, s_k) to PA(a_i)] on action option 2, the same process can be simulated to demonstrate the effects of conscious cognitive control as a top–down effect. Furthermore, it is also possible to combine both of these effects in a simulation.

4.8 Scenario 7: mismatch between the predicted and actual effects of an action

In most of the previous scenarios, the predicted effect always has positively affected action execution. Nevertheless, it may not always be the case and as pointed out in Sect. 1 (with the example of learning how to ride a bicycle) there may be a difference between what is predicted and what actually occurs. In this scenario, this phenomena will be simulated. For this simulation, a single input is considered but which triggers preparation of an action a_1 for which two options for an effect are considered: the first is b_1 which is the predicted effect of a_1, whereas the second b_2 is not predicted. The second option indicates what actually will occur after execution of a_1 (see Table 4). In Table 4, the weight changes for this case have been highlighted for each effect option separately. For the effect option 1 weight ω_3 [i.e. from EA(a_1) to WS(b_1)] was set '0' as the effect of actual execution is not b_1 and the same weight for effect option 2 was set '0.5' to facilitate the different non-predicted effect of the action (the same explanation for ω_4). The weight from WS(b_2) to SS(b_2) (i.e. ω_7 for option 2) was set 0.5 to facilitate the effects of actual sensing (this value can be further increased but 0.5 was selected merely to highlight an average effect in the sensing). Also according to the formation of this scenario at the beginning, the agent will only predict the effect b_1 of
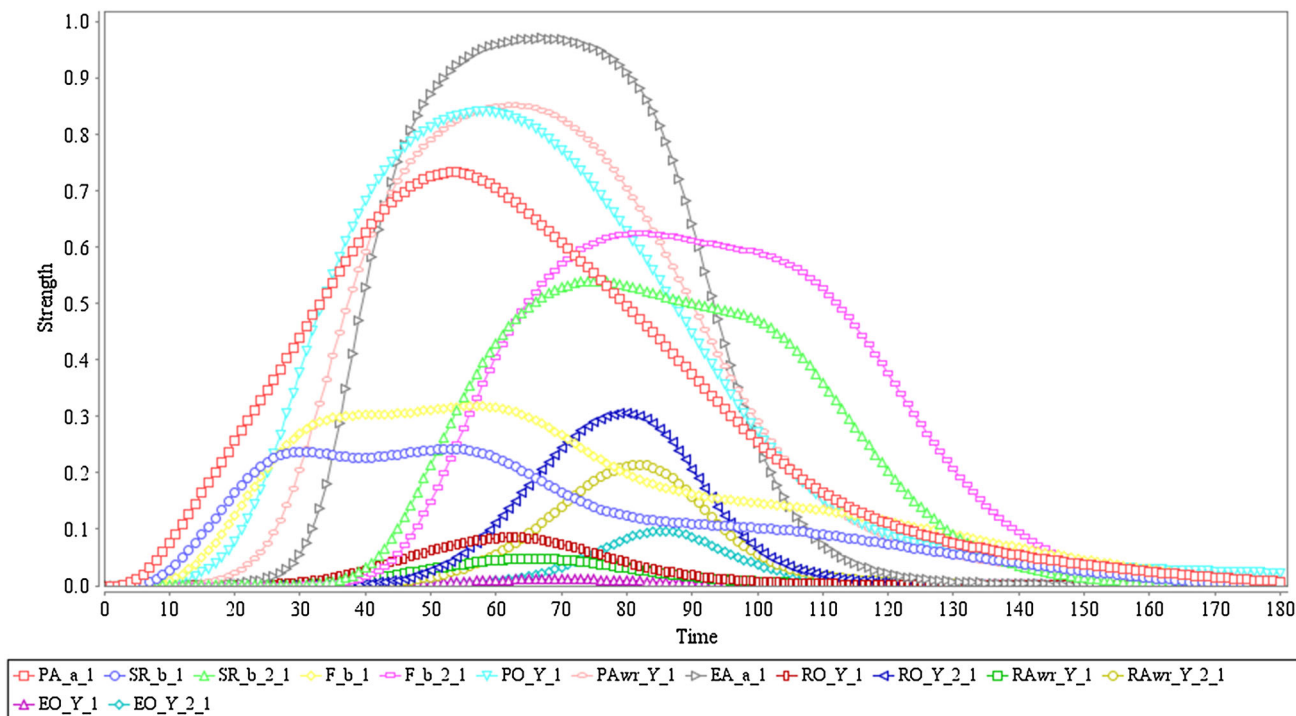
Fig. 8 Scenario 7: Mismatch between the predicted and actual effects of an action. In here 'Y_1' represents 'a_1, b_1, c_1, s_1' whereas 'Y_2' represents 'a_2, b_2, c_2, s_2'

action a_1, but not b_2. Therefore, the weight ω_{10} [i.e. PA(a_1) to SR(b_2)] was set '0' (the same explanation applies for weight ω_{12}). The weight ω_{30} from EA(a_1) to RO(a_1, b_1, c_1, s_1) was set '0' as there will not be any retrospective ownership for action a_1. Furthermore, the weight ω_{32} for effect option 2 [i.e. from PO(a_1, b_2, c_1, s_1) to RO(a_1, b_2, c_1, s_1)] was set '0' as there is no prior ownership on action option a_2. Additionally, the link from EA(a_1) to RAwr(a_1, b_1, c_1, s_1) (i.e. ω_{34}) was also set '0'. Also as there is no prior awareness related with effect option b_2, weight ω_{37} [i.e. from PAwr(a_1, b_2, c_1, s_1) to RAwr(a_1, b_2, c_1, s_1)] was also set '0'. The following execution trace is expected from the agent:

– External stimulus s_1 and context c_1 occur and trigger preparation of action a_1.
– Based on the preparation state for a_1, the sensory representation of a (positive) predicted effect b_1 of a_1 is generated.
– Based on this positive predicted effect and the other states, a prior ownership state for action a_1 is generated.
– Prior ownership for action a_1 is followed by prior awareness, which is generated just before the action execution.
– Prior ownership and prior awareness states for action a_1 lead to actual execution of action a_1.

– The execution of a_1 does actually not affect b_1 but a different effect b_2. Therefore, through sensing the agent will develop a sensory representation of b_2 and the feeling of b_2.
– Based on the generated states, after the execution of action a_1 the agent may develop a low retrospective ownership state for action a_1 with effect b_2 due to the conflict between predicted effect and sensed actual effect.
– Retrospective ownership of action a_1 with effect b_2 is followed by retrospective awareness of action a_1 with effect b_2.
– Finally the agent may not properly communicate this ownership and awareness (depend on the context).

The simulation of this scenario is shown in Fig. 8. As expected the agent triggers preparation of action a_1 at time point 3 (with the peak value of 0.73). Based on that the sensory representation of predicted effect b_1 of a_1 is generated (through the as-if body loop with peak value 0.23), followed by the feeling of b_1 (through the as-if body loop with the peak value 0.30). Next these states contribute to generate a prior ownership (for a_1 with effect b_1) which starts to trigger at the time point 5 with peak value 0.84. After activating the prior ownership, prior awareness develops, mainly upon the formation process of effect prediction b_1 for a_1. The prior awareness starts to pop up

around time point 12 and obtains peak value 0.85. As a result of the prior awareness and ownership states, the agent initiates the actual execution of action a_1 which propagates its actual effects through the (external) body loop. The action execution starts at time point 13 and its maximum strength is 0.97. In Fig. 8, it is clearly shown that the execution of action a_1 does not affect b_1 via sensing. Instead it has a different actual effect b_2; through sensing it affects the sensory representation and the feeling of b_2. Therefore, the sensory representation b_1 of a_1 does not behave as expected by adding the sensed actual effect to the predicted effect and the same effect does not propagate to the feeling of b_1 (cf. [20, 47, 48]). Two activations: that do occur are of the sensory representation of b_2 and feeling for b_2. They emerge just at that point in time as they were not there at the beginning. These new states provide satisfactory activation levels but contribute for relatively low retrospective ownership and awareness states (mainly due to the nonexistence of influences of prior ownership and awareness states, respectively). Subsequently due to these poor retrospective ownership and awareness states, agent does not properly communicate the action and its effect. Poor communication was noted experimentally in [62] through a card game together with a covert exchange to facilitate the conflict between predicted and actual outcome. Therefore, also from experimental findings this poor communication can be justified. Furthermore, in simulation scenario 1 when there is no conflict between predicted and occurring effect, prior awareness on sensed effect has initiated at time point 35, whereas in this simulation it occurs around time point 50. Therefore, the temporal gap between the action and its perceived sensory outcome is less when the awareness is pre-existing but it is more when the (prior) awareness is not involved, which is observable from these simulation results; this is referred as intentional binding [15].

This behaviour has shown the effects when there is a conflict between what predicted versus actual. In future work, this will be especially useful when learning should incorporate with the model in a dynamical form. Having this conflict between sufficiently large prior awareness and a very low retrospective awareness can be further explored with more neuroscientific evidences. This is useful for adaptive behaviours especially for situation awareness-driven applications.

4.9 Scenario 8: cognitive impairment for shifting to a positive feelings

In real life, agents will experience both positive and negative situations. Nevertheless, if an agent always sticks to negative thoughts while suppressing the possible positive thoughts, in the long run this will lead to a depression. In this scenario, the agent is preparing for two action options where one leads to a positive feeling $F(b_1)$, while the other leads to a negative feeling $F(b_2)$. According to the bias to negative feelings, the agent tends to select the negative action though both options share exactly the same weight values for all weights on each option, except for ω_{29} which is for prior awareness (which introduces a conscious intervention on action execution). For the weight ω_{29} on the first (positive) action option 0.2 was used and for the second (negative) option 0.8. Except this change all the other parameter values are identical and through this only the impact from biased conscious awareness towards a particular action is modelled. If the agent is executing the action which is associated to negative feeling (though at the same time a possible positive action also exists) this will affect the agent's mood (the mood of the agent has not been in the scope of this model but will be a future work). Having a negative mood all the time together with this biased awareness towards such negative actions will take the agent to a cognitive disorder state called depression. A healthy agent will have the ability to shift in a relatively short time period again to positive actions by correcting the biased influence. Therefore, this model can be used to simulate a depression situation also, but with further improvements in the model. For the biased cognitive impairment on negative actions through awareness, the following trace is expected:

– The external stimulus s_1 and context c_1 will occur and trigger preparation of actions a_1 and a_2.
– Based on the preparation state for a_1, the sensory representation of predicted effect b_1 of a_1 is generated and in the same way for a_2, the sensory representation of predicted effect b_2 of a_2 is generated.
– It is assumed to be that both effects positively contribute to the action formation process while the effect b_1 has a positive associated feeling, and the effect b_2 has a negative associated feeling.
– Based on this positive predicted effect and the other states, a prior ownership state for action a_1 and a_2 is independently generated.
– Prior ownership for action a_1 is followed by the prior awareness of a_1 and the prior ownership for action a_2 is followed by the prior awareness of a_2.
– Nevertheless, the prior awareness of a_2 will dominate the action formation process due to the biased competition.
– Due to this the states related to action a_1 will loose their activation, while the sates related to action a_2 will continue as the selected options.
– Prior ownership and prior awareness states for action a_2 lead to actual execution of action a_2.
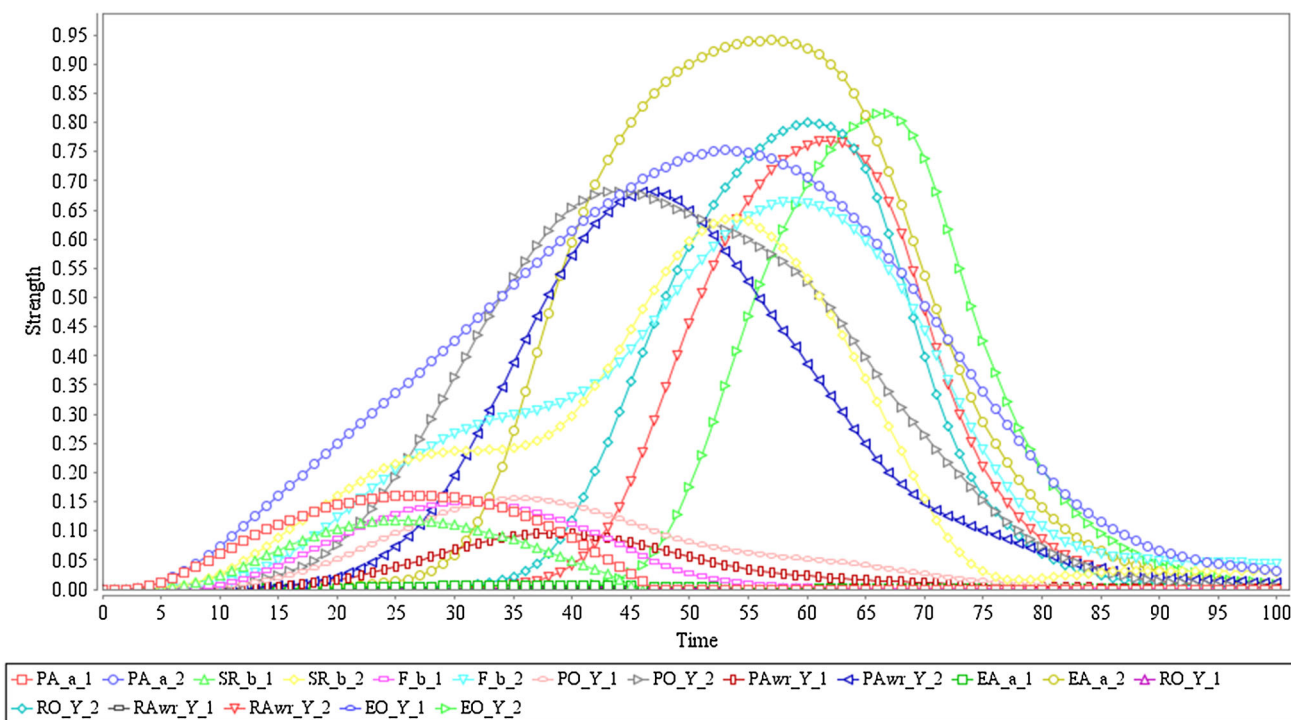
Fig. 9 Scenario 8: Cognitive impairment for shifting to a positive feeling. In here 'Y_1' represents 'a_1, b_1, c_1, s_1' whereas 'Y_2' represents 'a_2, b_2, c_2, s_2'

– The execution of a_2 affects b_2 in a positive manner through the sensing.
– The agent develops sensory representation of b_2 and the feeling of b_2 in line with what is predicted.
– Based on the generated states, the agent develops a retrospective ownership state for action a_2 with effect b_2.
– Retrospective ownership of action a_2 with sensed effect b_2 is followed by retrospective awareness of action a_2 with effect b_2.
– Finally the agent communicates this ownership and awareness.

The simulation of this scenario is shown in Fig. 9; it is in line with the expected trace. The agent has initiated two action preparations once the input stimuli are received. Both $PA(a_1)$ and $PA(a_2)$ are activated at time point 3 and almost with the same activation speed. Parallel to this the agent also initiates the sensory representation and feelings for effects b_1 and b_2 for actions a_1 and a_2, respectively. Furthermore, the rates of activation for these four states are also almost the same at the very beginning (until prior awareness states pop up). Followed by these states, prior ownership states are also activated for both options and then subsequently the prior awareness states for each option. Roughly at time point 10, the prior awareness states start to emerge and from that very moment the states related with the option 1 show decline effects. For example, the action preparation states initially have same speed but

after the time point 10 state $PA(a_1)$ is losing the speed, whereas $PA(a_2)$ maintains the same momentum and reaches maximum value 0.75. Similar to this $SR(b_1)$ reaches peak value 0.11 through the as-if body loop (i.e. through the predictive process), whereas $SR(b_2)$ reaches peak value 0.23 through the predictive process and 0.63 through the body loop. Furthermore, $F(b_1)$ reaches peak value 0.14 through the as-if body loop, whereas $F(b_2)$ reaches peak value 0.30 through the predictive process and 0.66 through the body loop. Moreover, $PO(a_1, b_1, c_1, s_1)$ reaches peak value 0.15, whereas $PO(a_2, b_2, c_2, s_2)$ reaches peak value 0.68. Also $PAwr(a_1, b_1, c_1, s_1)$ reaches only a 0.09 value but $PAwr(a_2, b_2, c_2, s_2)$ reaches 0.68.

These results clearly show the impact from conscious intervention for the action formation process (all the other unconscious related suppressive parameter values are identical for both options and therefore no bias was introduced by unconscious processes). Therefore, due to this bias the negatively associated action option a_2 has executed with peak value 0.92. Furthermore, it is shown that the execution of action a_2 (via the body loop) also affects in positive manner via sensing, the sensory representation of b_2 and the feeling of b_2. Due to these, the agent develops $RO(a_2, b_2, c_2, s_2)$ at time point 30 with peak value 0.79 and $RAwr(a_2, b_2, c_2, s_2)$ at the time point 34 with peak value 0.76. The agent does not develop any retrospective effects associated to action option a_1. Finally, the agent has

strongly communicated the ownership and awareness on action a_2.

5 Discussion and future work

Computational modelling is considered an important pillar for the development of cognitive science and related disciplines [63–65]. Moreover, the developments in brain imaging and recording techniques also strongly contribute to more and more focused phenomena explored in the cognitive, behavioural, affective and physiological research areas. Nevertheless, there is room for a more generic, compound process to explain a wide range of cognitive functionalities by aggregating many of these local but highly influential information from the computational perspective [66]. For example, Ron Sun in [66] has stated that "*integrative computational cognitive modelling may serve in the future as an antidote to the increasing specialization of scientific research*" [66], p. 14. On the other hand, there are many implications of these hypotheses (or evidences) and it would be beneficial if there was a mechanism that can be used to scrutinize these ideas or hypotheses using this as a workbench in much more abstract and global level (cf. [64]). Additionally, the human brain and its phenomena are immeasurably complex systems/processes that involve many uncountable factors that make experiments not always coherent with reality. Nevertheless, having computational models enables to uplift the progress of understanding these processes in a broader level as a multidisciplinary approach (cf. [63]).

From the Artificial Intelligence perspective, more and more complex systems related problems are addressed that include human cognitive aspects. For example, situations related to air traffic management, stock market analysis, business processes, human awareness of energy usage, cognitive impairments and various medical disorders, how a certain form of therapy can have its effect on a patient, etc. In many such situations, it is not practical to create real experimental setups to analyse the emergence of problems. Nevertheless, the importance and significance of such analysis is essential from a safety, performance and health perspective. Multi-agent-based simulation approaches have been noted as the potential solution for this [67] by considering agent-based simulations. Such simulations can capture emergent phenomena, providing a natural description of a system, and its flexibility: essentials for such complex system research. Even though agent-based simulations are a promising technology for this, there are many problems associated with it in terms of more realistic models for such agents to behave or perform. Often non nature inspired, simple heuristically determined rule-driven agent simulations are used for this, although reality is more

complicated and far away from those simplifications. It needs more realistic representations and analysis in much closer to the natural situation (cf. [67]). Therefore, in agent systems more realistic computational models that make use of the latest neuro-cognitive findings can be used to simulate agent behaviour in a more realistic manner. Especially when it comes to the systems that include human cognition factors in dynamic systems, nature inspired cognitive computational models have more power to provide realistic results. Therefore, computational cognitive modelling as a multidisciplinary approach has many benefits for both cognitive science and artificial intelligence.

This paper presented a computational cognitive model for action awareness focusing on action preparation and performance by considering its cognitive effects and affects from both prior and retrospective perspective relative to the action execution. It is a fundamental research question what is the human cognition behind the action selection and how this is related to conscious and unconscious elements. Mainly for two hypotheses or claims, attention has been obtained by the community, as presented in Sects. 1 and 2. However, it is not yet clear what is the exact process behind this human cognition. The first hypothesis (by Benjamin Libet and others) claimed that humans may prepare for and perform actions without being conscious of these preparation and execution processes, and the awareness of motor intention of this action is not causing the behaviour, but comes after the action preparation and relatively just before the action execution time [4–8]. The second hypothesis (by Haggard and co-workers) claims that awareness of an action is a dynamic combination of both prior awareness and retrospective awareness through predictive motor control and inferential sense-making relative to the action execution, respectively [20, 21]. Furthermore, the intentional binding effect [15] shows the impact of awareness on action selection and this contributes to the second hypothesis to bring out the influence of awareness on action selection. Although these two hypothesises seem to contradict each other from the semantic point of view, by other research (cf. [5]) from a pragmatic point of view it seems hard to generally reject any of these claims on the basis of the available empirical evidences. This paper utilizes both ideas into a compound process (together with other supportive processes) and scrutinizes the behaviour through related scenarios. This work was not conducted from scratch, but adopts parts of the model presented in [23], mainly for the mechanisms of action ownership and other unconscious states/processes (e.g. action preparation, sensory representation, effect prediction process, mirroring, etc.). Having that previous model which was validated through simulations mainly for unconscious action formation, in this paper its scope is further extended to incorporate the conscious aspects related to action selection. The main research questions for this work are

(1) How does the internal prediction process shape or contribute to the (prior) awareness of the action?

(2) How does the inferential sense-making shape or contribute to the (retrospective) awareness of the action execution?

(3) How does the awareness contribute to action execution?

(4) What is the relation and interplay between conscious and unconscious action formation through action ownership and relevant awareness states?

Each research question is explored from the cognitive science perspective and analysed through the modelling perspective (with simulations) to isolate a working definition. For the internal prediction process, the hypothesis of Damasio's as-if body loop is used as a basis, as previously presented in [23]. In the current model, it is further extended by embedding an unconscious process referred as the lateral inhibition (see [49]) to strengthen the competition among action options through the as-if body loop. With that new addition the unconscious action prediction process is coupled with the prior awareness state (as a higher order cognitive state) to facilitate the conscious aspects as proposed by Haggard et al.: predictive processes are also playing a role in action awareness (e.g. see [20, 21]). Also given empirical evidence to support the idea that awareness of motor intention for an action comes after the action preparation and relatively just before the action execution time (cf. [4–8]), in this model it is ensured that the awareness states are always higher order cognitive states which are not getting affected by most of the low level cognitive states. Predictive processes anticipate effects of each action option and lead to a competition to get selected a GO signal. The basic decision-making is assumed to be based on a feeling-related valuations associated with the effects of each action option. These internally predicted feelings work as a bottom–up feedback to develop a coherent conscious experience of action selection, mainly on action options that have strong predictive effects; therefore, actions with poor predictive effects do not get any conscious attention. Therefore, prior awareness state is only affected by the feeling and prior ownership states. The simulation results have confirmed that when there is a prior awareness it always appears just before the action execution.

The same approach is applied to the second research question, also to isolate a working definition. Once the feedback sensory information on the effects of the actual action execution is available and is integrated with the predicted effects as suggested in [20, 47, 48] to evaluate the binding of what is predicted and the actual effect. Through this sense-making process, the agent will experience the retrospective effects as presented in previous work [23],

mainly for acknowledging authorship of an action, reflection on one's own functioning, personal learning and development. This was further extended with a new abstract cognitive state name called retrospective awareness which is responsible for more conscious interpretation of retrospective effects of the action execution. Having empirical evidences for the contribution of inferential processes on action awareness from Haggard and co-workers' experiments (cf. [20, 21]), this extension can be justified from the cognitive neuroscience perspective. Furthermore, through simulations on relevant scenarios this extension was validated. Especially the seventh scenario clearly shows that when there is a mismatch between what is predicted and what actually occurs, the two steps sigmoid behaviour on the sensory representation and the feeling states is not observed, whereas a poor retrospective awareness on the effect was predicted, as highlighted in [20]. When considering the first scenario and the seventh, this may even contribute for the observation of intentional binding (cf. [15]) mainly through the feeling state: increasing the same feeling through inferential information and forming a completely new feeling through the inferential process. In the seventh simulation, the agent does not aggregate the prior awareness to the retrospective awareness and therefore the agent will not be consciously able to communicate the conflict or mismatch. This also shows the subjective time effect found in the intentional binding: when there is no conflict between predicted and occurring effect, the prior awareness on sensed effect has initiated earlier than when a conflict occurs. Therefore, the feeling of the temporal gap between the action and its perceived sensory outcome can be explained through the retrospective awareness [15]. To have these results in this model, the retrospective awareness states were mainly affected by four states: action execution of a_i, feeling of effect b_i, prior awareness for action a_i with b_i, c_k, and s_k and prior ownership for action a_i with b_i, c_k and s_k.

The third research question shows many thoughts within the cognitive neuroscience community. It suggests that awareness is not a cause for an action execution but it seems like an after-effect of a set of unconscious cognitive processes leading to the action [1, 4–8]. The other idea shows that an additional influence through awareness may inject some bias or effect on decision-making especially with the findings of intentional binding [15, 21, 37, 43]. There are empirical evidences to support the second claim (at least for some inputs), emphasising why it is hard to accept the first claim due to its non moderate statistical strength on empirical evidences: in general only $\sim 60\%$ accuracy in experimental data (see [5]). Therefore, this model includes both features: awareness is not required for action execution and awareness may play a role in action

execution, depending on the specific scenario. In other words, the model handles both the purely unconscious action formation and the hybrid form of action formation with both conscious and unconscious elements. Furthermore, through the simulation results also this was validated. Especially the first and second scenarios provide information to show that the agent is able to execute an action with and without awareness. The eighth scenario presents a situation in which biased awareness may transform a healthy person into a cognitively impaired position. By continuing the cognition behind the eighth scenario for a longer period of time, this can even explain the effects of a patient suffering from a depression (this will be considered as a future work). Therefore, with different settings and scenarios, the model is capable to demonstrate the contribution of awareness from zero to high. Therefore, the model facilitates a good spectrum to represent the effects of awareness on action selection. There are situations, for example like flight or fight situations, which mostly show the unconscious action formation side of the spectrum that includes very quick and strong action executions (cf. [68–70]). According to the second scenario, it is clear that when agent is purely performing in an unconscious mode, the strength of action execution is relatively weak. Therefore, it may be possible to add improvements to this model and its settings, especially on the unconscious perspective including the emotion related effects. Further information on the interplay between bottom–up and top–down processes seems to be useful for such further improvements of this model (see [68–71]).

The fourth research question is realised as an aggregation of the other three questions. The interplay between conscious and unconscious action formation is mainly realised through having effects from feeling and ownership states on the awareness states (unconscious to conscious), and effects from prior awareness states to the action preparation and action execution states (conscious to unconscious). The first links (bottom–up) play a role to pass low level information to develop awareness of what is going on in a high level form, whereas the second type of links (top–down) contributes to inject some bias or excitement to intentionally drive the action formation. When there is poor activation of the bottom–up links, agent is unable to develop prior awareness and therefore, as presented in the second scenario, the agent can perform the action selection in purely unconscious form by the feeling of ownership. When the agent is having sufficient activations in bottom–up links, awareness develops and shows how that leads to an action execution, as in the first scenario. Also particularly in the eighth scenario, the role of top–down links demonstrates the power of an intentional focus on action selection. The third simulation highlights why agents are not always performing a task even in

unconscious mode. This shows that even in unconscious form it is necessary to have a sufficiently large satisfactory predicted effect to have a GO signal. Therefore, this shows the mechanism of vetoing in unconscious form. The cognitive control is a useful process to explain the interplay between conscious and unconscious action formation. The fifth and sixth simulations present the role of cognitive control. In the fifth simulation, it shows how a particular option is getting suppressed by the other option through cognitive control just by having a value of 0.2 difference for the links between action preparation to sensory representation for two options. The sixth simulation shows that once the dominant option of the fifth simulation completed, the suppressed action option emerges due to not having the effects of cognitive control and, more importantly, with strong activations for each state. Through this, it is clear from the simulation perspective that the model can have different configurations to facilitate different behaviours both from the unconscious and the combination of conscious and unconscious forms. Also as mentioned in the other three research questions, the cognitive neuroscience basis behind this model was inspired by the experiments and evidences found in the literature.

Having interesting simulation results for many scenarios, still there is more to improve on this work. It was possible to isolate a generic parameter value set that worked for 8 simulations. A generic parameter value set provides more confidence from the model validation perspective and its practical usage in future complex applications. Nevertheless, it is not a trivial task to find such generic parameter value set. The approach used to identify this parameter value set was explained in Sect. 4.1, but it is not a fully automated process. Due to the complexity of the human brain and limitations in measuring techniques of human cognition, there are no detailed numerical empirical data to use in this process, but only some characteristics of behaviours are available in more fuzzy form. Due to this issue, it is not possible to directly use parameter estimation techniques available for dynamic systems. Therefore, it is useful to explore the parameter estimation techniques particular to the characteristics of these types of work. Furthermore, from the cognitive neuroscience perspective, this model has many more areas to explore both in conscious and unconscious levels. In the third scenario, it shows the vetoing process in unconscious form, but the same process with awareness which is referred as intentional inhibition is a future work for this model. Human awareness has its specializations, for example, emotional awareness, situation awareness. Working processes behind these concepts are more complicated and need further research to incorporate those into this model.

Finally, this model may be useful in many applications. Especially for agent-based simulations on complex systems

that need action selection related to cognitive aspects. Also, this model can be used as a basis for subsequent work in developing ambient agent systems able to monitor, analyse and support persons trying to develop a healthy lifestyle. If such systems have such a model of the underlying human processes, they can use this to have a deeper understanding of the human.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Custers R, Aarts H (2010) The unconscious will: how the pursuit of goals operates outside of conscious awareness. Science 329:47–50. doi:10.1126/science.1188595
2. Cohen JD, Schooler JW (1997) Scientific approaches to consciousness. L. Erlbaum Associates, Mahwah
3. Blackmore SJ (2010) Consciousness an introduction. Hodder Education, London
4. D'Ostilio K, Garraux G (2012) Brain mechanisms underlying automatic and unconscious control of motor action. Front Hum Neurosci. doi:10.3389/fnhum.2012.00265
5. Haynes J-D (2011) Decoding and predicting intentions: predicting intentions. Ann N Y Acad Sci 1224:9–21. doi:10.1111/j.1749-6632.2011.05994.x
6. Baumeister RF, Masicampo EJ, Vohs KD (2011) Do conscious thoughts cause behavior? Annu Rev Psychol 62:331–361. doi:10.1146/annurev.psych.093008.131126
7. Wegner DM (2002) The illusion of conscious will. MIT Press, Cambridge
8. Libet B, Gleason CA, Wright EW, Pearl DK (1983) Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. Brain 106:623–642. doi:10.1093/brain/106.3.623
9. Mesulam M (1998) From sensation to cognition. Brain 121:1013–1052. doi:10.1093/brain/121.6.1013
10. McIntosh AR (2000) Towards a network theory of cognition. Neural Netw 13:861–870. doi:10.1016/S0893-6080(00)00059-9
11. Swanson LW (2012) Brain architecture: understanding the basic plan, 2nd edn. Oxford University Press, New York
12. Fries P (2005) A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn Sci 9:474–480. doi:10.1016/j.tics.2005.08.011
13. Bressler SL (1995) Large-scale cortical networks and cognition. Brain Res Rev 20:288–304. doi:10.1016/0165-0173(94)00016-I
14. Friston K (2002) Beyond phrenology: what can neuroimaging tell us about distributed circuitry? Annu Rev Neurosci 25:221–250. doi:10.1146/annurev.neuro.25.112701.142846
15. Haggard P, Clark S, Kalogeras J (2002) Voluntary action and conscious awareness. Nat Neurosci 5:382–385. doi:10.1038/nn827
16. Moore JW, Lagnado D, Deal DC, Haggard P (2009) Feelings of control: contingency determines experience of action. Cognition 110:279–283. doi:10.1016/j.cognition.2008.11.006
17. Walsh E, Haggard P (2013) Action, prediction, and temporal awareness. Acta Psychol (Amst) 142:220–229. doi:10.1016/j.actpsy.2012.11.014
18. Desantis A, Hughes G, Waszak F (2012) intentional binding is driven by the mere presence of an action and not by motor prediction. PLoS ONE 7:e29557. doi:10.1371/journal.pone.0029557
19. Moore JW, Obhi SS (2012) Intentional binding and the sense of agency: a review. Conscious Cogn 21:546–561. doi:10.1016/j.concog.2011.12.002
20. Moore J, Haggard P (2008) Awareness of action: inference and prediction. Conscious Cogn 17:136–144. doi:10.1016/j.concog.2006.12.004
21. Haggard P, Clark S (2003) Intentional action: conscious experience and neural prediction. Conscious Cogn 12:695–707. doi:10.1016/S1053-8100(03)00052-7
22. Thilakarathne DJ, Treur J (2013) Modelling prior and retrospective awareness of actions. In: Ferrández Vicente JM, Álvarez Sánchez JR, Paz López F, Toledo Moreo FJ (eds) Natural and artificial models in computation and biology. Springer, Berlin, pp 62–73
23. Treur J (2012) A computational agent model incorporating prior and retrospective ownership states for actions. Biol Inspired Cogn Archit 2:54–67. doi:10.1016/j.bica.2012.06.001
24. Decety J, Sommerville JA (2003) Shared representations between self and other: a social cognitive neuroscience view. Trends Cogn Sci 7:527–533. doi:10.1016/j.tics.2003.10.004
25. Chaminade T, Marchant JL, Kilner J, Frith CD (2012) An fMRI study of joint action–varying levels of cooperation correlates with activity in control networks. Front Hum Neurosci. doi:10.3389/fnhum.2012.00179
26. Iacoboni M (2009) Mirroring people: the science of empathy and how we connect with others. Picador, New York
27. Iacoboni M (2008) Mesial frontal cortex and super mirror neurons. Behav Brain Sci. doi:10.1017/S0140525X07003214
28. Iacoboni M, Dapretto M (2006) The mirror neuron system and the consequences of its dysfunction. Nat Rev Neurosci 7:942–951. doi:10.1038/nrn2024
29. Blakemore S-J, Wolpert D, Frith C (2000) Why can't you tickle yourself? NeuroReport 11:R11–R16. doi:10.1097/00001756-200008030-00002
30. Wolpert DM, Kawato M (1998) Multiple paired forward and inverse models for motor control. Neural Netw 11:1317–1329. doi:10.1016/S0893-6080(98)00066-5
31. Voss M, Ingram JN, Haggard P, Wolpert DM (2006) Sensorimotor attenuation by central motor command signals in the absence of movement. Nat Neurosci 9:26–27. doi:10.1038/nn1592
32. Kilner JM, Friston KJ, Frith CD (2007) Predictive coding: an account of the mirror neuron system. Cogn Process 8:159–166. doi:10.1007/s10339-007-0170-2
33. Blakemore S-J, Wolpert DM, Frith CD (1998) Central cancellation of self-produced tickle sensation. Nat Neurosci 1:635–640. doi:10.1038/2870
34. Eickhoff SB (2006) The Human Parietal Operculum. I. Cytoarchitectonic Mapping of Subdivisions. Cereb Cortex 16:254–267. doi:10.1093/cercor/bhi105
35. Inoue K (2002) Role of human SII cortices in sensorimotor integration. Clin Neurophysiol 113:1573–1578. doi:10.1016/S1388-2457(02)00162-1
36. Blakemore S-J, Frith C (2003) Self-awareness and action. Curr Opin Neurobiol 13:219–224. doi:10.1016/S0959-4388(03)00043-6
37. Haggard P (2005) Conscious intention and motor cognition. Trends Cogn Sci 9:290–295. doi:10.1016/j.tics.2005.04.012
38. Frith C (2000) Explaining the symptoms of schizophrenia: abnormalities in the awareness of action. Brain Res Rev 31:357–363. doi:10.1016/S0165-0173(99)00052-1
39. Marcel AJ (2003) The sense of agency: awareness and ownership of actions and intentions, Agency Self-Aware. Issues Philos. Psychol. Oxford University Press, Oxford
40. Monsell S (2003) Task switching. Trends Cogn Sci 7:134–140. doi:10.1016/S1364-6613(03)00028-7
41. Banks WP, Pockett S (2007) Benjamin Libet's work on the neuroscience of free will. In: Velmans M, Schneider S (eds)

Blackwell companion conscious. Blackwell Publishing, Malden, pp 657–670

42. Wolpert DM (1997) Computational approaches to motor control. Trends Cogn Sci 1:209–216. doi:10.1016/S1364-6613(97)01070-X

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

44. Damasio AR (2005) Descartes' error: emotion, reason, and the human brain. Penguin, London

45. Damasio AR (2000) The feeling of what happens: body and emotion in the making of consciousness. Harcourt Inc., New York

46. Damasio AR (2012) Self comes to mind: constructing the conscious brain. Vintage Books, New York

47. Synofzik M, Thier P, Leube DT, Schlotterbeck P, Lindner A (2010) Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions. Brain 133:262–271. doi:10.1093/brain/awp291

48. Voss M, Moore J, Hauser M, Gallinat J, Heinz A, Haggard P (2010) Altered awareness of action in schizophrenia: a specific deficit in predicting action consequences. Brain 133:3104–3112. doi:10.1093/brain/awq152

49. Picard N, Strick PL (2001) Imaging the premotor areas. Curr Opin Neurobiol 11:663–672. doi:10.1016/S0959-4388(01)00266-5

50. Fried I, Katz A, McCarthy G, Sass KJ, Williamson P, Spencer SS, Spencer DD (1991) Functional organization of human supplementary motor cortex studied by electrical stimulation. J Neurosci 11:3656–3666

51. Aron AR (2007) The neural basis of inhibition in cognitive control. Neuroscientist 13:214–228. doi:10.1177/1073858407299288

52. Blakemore S-J, Frith CD, Wolpert DM (1999) Spatio-temporal prediction modulates the perception of self-produced stimuli. J Cogn Neurosci 11:551–559. doi:10.1162/089892999563607

53. Fourneret P, de Vignemont F, Franck N, Slachevsky A, Dubois B, Jeannerod M (2002) Perception of self-generated action in schizophrenia. Cogn Neuropsychiatry 7:139–156. doi:10.1080/13546800143000212

54. Frith CD (1992) The cognitive neuropsychology of schizophrenia. L. Erlbaum Associates, Hove

55. Treur J (2013) An integrative dynamical systems perspective on emotions. Biol Inspired Cogn Archit 4:27–40. doi:10.1016/j.bica.2012.07.005

56. James W (1884) What is an emotion? Mind 9:188–205. doi:10.1093/mind/os-IX.34.188

57. Bosse T, Jonker CM, Van Der Meij L, Treur J (2007) A language and environment for analysis of dynamics by simulation. Int J Artif Intell Tools 16:435–464. doi:10.1142/S0218213007003357

58. Thilakarathne D (2014) Neurologically inspired computational cognitive modelling of situation awareness. In: Ślęzak D, Tan A-H, Peters J, Schwabe L (eds) Brain Informatics Health. Springer International Publishing, Poland, pp 459–470

59. Lewandowsky S (2011) Computational modeling in cognition: principles and practice. Sage Publications, Thousand Oaks

60. Aleman A, Kahn R (2005) Strange feelings: do amygdala abnormalities dysregulate the emotional brain in schizophrenia? Prog Neurobiol. doi:10.1016/j.pneurobio.2005.11.005

61. Kring AM, Caponigro JM (2010) Emotion in schizophrenia: where feeling meets thinking. Curr Dir Psychol Sci 19:255–259. doi:10.1177/0963721410377599

62. Johansson P, Hall L, Sikström S, Olsson A (2005) Failure to detect mismatches between intention and outcome in a simple decision task. Science 310:116–119. doi:10.1126/science.1111709

63. Shiffrin RM (2010) Perspectives on modeling in cognitive science. Top Cogn Sci 2:736–750. doi:10.1111/j.1756-8765.2010.01092.x

64. McClelland JL (2009) The place of modeling in cognitive science. Top Cogn Sci 1:11–38. doi:10.1111/j.1756-8765.2008.01003.x

65. Addyman C, French RM (2012) Computational modeling in cognitive science: a manifesto for change. Top Cogn Sci 4:332–341. doi:10.1111/j.1756-8765.2012.01206.x

66. Sun R (2008) Introduction to computational cognitive modeling. In: Sun R (ed) The cambridge handbook of computational psychology. Cambridge University Press, Cambridge, pp 3–20

67. Bonabeau E (2002) Agent-based modeling: methods and techniques for simulating human systems. Proc Natl Acad Sci 99:7280–7287. doi:10.1073/pnas.082080899

68. Carretié L, Hinojosa JA, Martín-Loeches M, Mercado F, Tapia M (2004) Automatic attention to emotional stimuli: neural correlates. Hum Brain Mapp 22:290–299. doi:10.1002/hbm.20037

69. McRae K, Misra S, Prasad AK, Pereira SC, Gross JJ (2012) Bottom–up and top–down emotion generation: implications for emotion regulation. Soc Cogn Affect Neurosci 7:253–262. doi:10.1093/scan/nsq103

70. Ochsner KN, Ray RR, Hughes B, McRae K, Cooper JC, Weber J, Gabrieli JDE, Gross JJ (2009) Bottom–up and top–down processes in emotion generation: common and distinct neural mechanisms. Psychol Sci 20:1322–1331. doi:10.1111/j.1467-9280.2009.02459.x

71. Vuilleumier P, Armony J, Clarke K, Husain M, Driver J, Dolan R (2002) Neural response to emotional faces with and without awareness: event-related fMRI in a parietal patient with visual extinction and spatial neglect. Neuropsychologia 40:2156–2166. doi:10.1016/S0028-3932(02)00045-3

Dilhan J. Thilakarathne is a Ph.D. student in the Agent Systems Research Group of the Department of Computer Science at the VU University Amsterdam. His educational background includes Artificial Intelligence and Computer Science. His primary research is focused into analysing the findings of cognitive/affective/behavioural sciences on processes of human cognition and behaviour, and designing computational cognitive models that are inspired from those. Special attention is given to both unconscious and conscious processes and their interactions. Some of the cognitive phenomena that have been explored are: conscious and unconscious action selection, intentional inhibition, emotional awareness, situational awareness, perceptual load, and cognitive controlling.

Jan Treur has a full professorship in Artificial Intelligence at VU University Amsterdam since 1990, and leads the Agent Systems Research Group. He is an internationally well-recognized expert in agent-oriented, cognitive and social modelling. His research area is highly multidisciplinary and takes into account knowledge from Cognitive Sciences, Social Sciences and Cognitive, Affective and Social Neuroscience, which is used as a basis for scientifically justifiable computational models for human processes. Such computational models are used as a basis for socially aware computing applications: smart applications and software agents used to monitor, analyse and support humans in various aspects of their processes and lifestyle. Moreover, he initiated and designed a successful strongly multidisciplinary Bachelor and Master study program at VU University in Lifestyle Informatics and Socially Aware Computing, combining subjects from Ambient and Artificial Intelligence, Computer Science, Biomedical Sciences, Neurosciences, Psychology, and Social Sciences.