

A Model of Agential Learning Using Active Inference

Riddhi J. Pitliya^{1,2} and Robin A. Murphy²

¹ VERSES Research Lab, Los Angeles, California, 90016, USA

² Department of Experimental Psychology, University of Oxford, Oxford, UK

Abstract. Agential learning refers to the process of forming beliefs regarding one’s degree of control over actions and outcomes in their environment. We first provide an overview and evaluation of associative, statistical, and Bayesian models of agential learning. We then argue that the existing models have limitations in explaining the process of agential learning. Finally, we introduce an active inference account of agential learning, and present results from simulations. We propose that the active inference framework may provide a comprehensive model of agential learning describing three fundamental processes: (i) perception, (ii) learning, and (iii) action.

Keywords: Agency · Agential Learning · Active Inference · Computational Psychology

1 Introduction

An agent is *someone* or *something* that acts to control their actions and events in the environment. Agency, then, refers to having control over one’s own actions, and leveraging that sense to control themselves or events in the environment [1], [2]. Agential learning is the process of tracking and forming relevant beliefs [3] regarding one’s degree of agency. Having an ongoing registration of the degree of control agents (self and others) have over the states in their environment facilitates individual- and group-level goal-directed behaviours [4].

Rather than a binary concept, degree of agency refers to the amount of control the agent has to generate or prevent the occurrence of the event. When based purely on objective experience, agency can be formalised as a statistical relationship, or contingency, between the action produced by an agent and its consequence/outcome (discrete variables), each with a dichotomous state of being present or absent. Contingencies, and correlations, vary on a scale from -1 to +1: a positive contingency is when an action predicts the outcome (e.g., pressing a button and the light being turned on), negative contingency is when an action signals the absence of the outcome (e.g., pressing a button and the light being turned off), and zero contingency is when the action has no relation to the presence or absence of the outcome (e.g. the pressing of a button does not have an impact on the light).

One experimental task widely used to assess action-outcome contingency learning involves an action that the participant can freely perform (a so-called

free-operant procedure [5]), such as pressing a button, and depending on the objective contingency set by the experimenter, an outcome is present or absent, such as a light being on or off. Subsequently, participants report the degree of control they perceive they have on a visual analog or numeric rating scale varying from -1 to +1. It has been well-demonstrated that perceived contingency as reported on the rating scale is aligned with the action-outcome objective contingency [6]–[12]. In this paper, we examine agential learning in the context of a simple scenario involving a single action and single outcome, though more complex versions could be entertained.

2 Previous Models of Agential Learning

Philosophers and then psychologists have been challenged to explain how agents learn that one event predicts (or causes) the presence or absence of another event [13]–[17]. Models that were originally used to explain cue-outcome contingency learning in non-human animals have been employed to explain human performance with some success. A number of models based on associative learning theory, statistical accounts, inferential reasoning, and Bayesian learning have been proposed. However, none comprehensively account for the complexity of learning [18]–[20].

2.1 Associative Models

Associative models [21]–[25] adopt a bottom-up approach and are process-driven. One model first applied to Pavlovian learning and then extended to explain instrumental learning is the Rescorla-Wagner model. Based on reinforcement principles and successfully applied to human statistical learning, it has a learning rule which is as follows:

$$\Delta V_n = \alpha\beta(\lambda_n - V_{total}) \quad (1)$$

ΔV represents the change in associative strength of a cue in that trial (n). The learning rate parameters, α and β , represent the associability of the cue and outcome respectively, representing how fast a particular action can be learnt. The subtraction in the parenthesis represents the prediction error, which is the discrepancy between the expected and actual occurrence of the outcome given a stimulus or action. λ represents the absolute value of the outcome on a trial (n). V_{total} is the total current associative prediction of all stimuli presented at that trial, therefore it comprises $V_1 + V_2 + \dots + V_n$. In sum, the Rescorla-Wagner model proposes that learning involves forming associations between all stimuli present in the environment, and those associations compete with one another as there is a limit to the amount of associative strength the outcome can support. The agent’s knowledge regarding associations is represented as a single weight value on each cue.

Associative learning explanations of action-outcome contingency learning suggest that agents integrate information online as each cue’s associative strength gets updated, requiring few cognitive resources and being computationally cheap. It is often described as a form of model-free associative learning. However, because the values get updated with each trial, explaining phenomena such as

retrospective revaluation, which has been demonstrated in humans [26]–[28], require additional assumptions.

The statistical account of action-outcome contingency learning [13], [29] suggests that the perceived contingency by the agent is related to an estimate of the difference between the probability of outcome occurring given an action and probability of outcome occurring given a lack of action. Models based on such statistical metrics alone, however, are unable to account for learning curves because probabilities are not affected by the amount of evidence on which they are based [30].

The associative learning and statistical models explain agential learning as the perception of an punctate value reflecting the action-outcome contingency, overlooking other processes that may be mechanistically involved. For example, the actions of the agent are not accounted for, except in the obvious cases where an experimenter impels or instructs action. An agent’s actions produces data for the agent, which they would use to form beliefs about their agency [31]. Indeed, a link between probability of acting and objective contingency has been established in free-operant tasks [32]–[34]. Moreover, agency may emerge from a form of inferential reasoning [18], [35], [36], wherein agents not only rely on direct sensory input and statistical metrics, but also engage in processes that involve learning about the dynamics and causes of the latent states of the world. Bayesian models of contingency learning provide an alternate account and address some of the limitations of previous models [37], [38].

2.2 Bayesian Associative Models

Inferring a Causal Structure

Researchers have proposed that agents may conduct Bayesian inference to infer the causal structure of the environment [39], [40]. The agent may do this by using bottom-up sensory information (observations) to infer the causal hidden states of their environment using an internal model of the world: a generative model that captures the agent’s beliefs about how (potentially dynamic) latent states of the world relate to observable sensory data.

A generative model of agential learning would comprise: (i) a prior probability distribution which represents the agent’s current beliefs about the hidden states, and (ii) a likelihood probability distribution which captures the agent’s knowledge of how observations (the action and outcome) are generated from hidden states by encoding the likelihood of observations given states. Using Bayes’ rule, one can compute a posterior probability distribution over hidden states, given observations. This can be interpreted as the agent’s beliefs regarding which hidden states best explain its sensory data, i.e., beliefs regarding their degree of agency. In the context of Bayesian cognitive neuroscience, this updating of beliefs via Bayesian inference has been analogised to perception [41].

The discrepancy between the agent’s predictions (from the priors) and beliefs about hidden states after receiving observations (posterior) is quantified by Bayesian surprise, a similar metric to prediction error as in the Rescorla-Wagner model. This is a measure of the degree to which the internal model and posterior beliefs get updated to reduce future surprise, which would ensure an internal

model of the causal structure of the world to be as close to the real causal structure of the world as possible. Bayesian inference can therefore be framed as an alternative problem of maximising marginal (log) likelihood, or, in other words, minimising surprise.

In traditional models of contingency learning, punctate values represent all of the agent’s knowledge. Bayesian approaches assume a different knowledge representation in the generative model as the agent entertains a probabilistic representation of its world, allowing a spectrum of alternative hypotheses to be represented via their posterior beliefs. The probability distributions allow the agent to express uncertainty, where the more spread out the beliefs are (represented by a flatter probability distribution), the greater the uncertainty. Such a representation of knowledge allows the model to keep track of multiple combinations of hypothetical beliefs, making the perceived causal structure malleable. Therefore, when belief regarding an association is highly uncertain, observational data has a rapid influence on changing that belief. These properties of a Bayesian approach account for how an agent perceives an action-outcome contingency [37].

Explaining Actions

The models described so far consider the agent as a passive observer, and predict action based on the action that is strongest associated with the outcome to produce the most favoured outcomes [7]. However, in reality, when agents are learning the degree of agency they have, the agent has the opportunity to explore or manipulate the world in order to extract information. In other words, the agent actively samples the environment, creating observations for itself to infer and perceive (a degree of) agency and test its beliefs in order to attain the preferred outcome state.

The representation of uncertainty in Bayesian models used to explain observational learning can be leveraged to guide active learning. Here, the agent’s actions are explained as the agent actively engaging with the environment to maximise expected information gain based on the generative model to reduce uncertainty [42]. While this explains exploratory behaviour, exploitation is explained by a separate function, based on Bayesian decision theory (or expected utility theory), wherein a value function of states is computed, which represents how rewarding the state is for the agent to be in. The value of the states depend on the agent’s learning history of state-action pairs, i.e., tracking how many times the agent attains the outcome by conducting that action from that state. The agent would thereby select the action that yields the outcomes it wants.

However, while exploitation and exploratory behaviour can be explained by different functions, the balance between the two often must be adjusted by introducing trade-off parameters, and different strategies have been employed according to task constraints [43]. This calls for a universal model of active learning instead of selecting a model from a class of models to optimally conduct the trade-off between exploration and exploitation dependent on the context. In the next section, we introduce an active inference model of agential learning, where a trade-off between information gain and rewards inherently arises as perception

and action are not treated as processes optimising two different functions but rather a single function. It is argued that the active inference framework provides a comprehensive model of agential learning.

3 Active Inference

3.1 Perception, Action, and Learning in Active Inference

Active inference is a process theory, based on the free energy principle [44], that provides a unified account of perception, action, and learning in agents. Active inference extends the (variational) Bayesian inferential process described earlier for perception to action, stemming from the notion that the agent minimises surprise. In active inference, however, a proxy for Bayesian surprise, (variational) free energy, is minimised [31], [45]. It is argued that while Bayesian frameworks consider surprise to be dependent on the agent’s generative model, surprise is also dependent on observations [45]. Active inference leverages this dependence to predict actions, wherein the agent infers the consequences of its own actions and the hidden states of the world, to exhibit behaviour that attains its preferences and actively reduces uncertainty in the agent’s world model [46], [47].

Under active inference, action selection is not only a function of past and present observations (as in Bayesian accounts), but also a function of prospective forms of inference based on anticipated future observations. The agent infers the best action sequence (policies) on the basis of future observations the actions would engender, which is based on beliefs about likelihoods of observations given the anticipated states in addition to the transitions of states across time as a function of the policy. This formulation of action selection in active inference casts action trajectories as a functional of beliefs (i.e., beliefs of beliefs, with probability distributions) inevitably encompassing the notions of uncertainty and preferences.

According to active inference, action selection occurs by expected free energy (EFE) being calculated for each policy and a policy is selected according to its negative EFE as policies that afford the lowest EFE are the most likely. EFE can be seen as the combination of (i) the anticipated information gain afforded by expected observations under a policy (exploration) and (ii) how well expected observations align with preferences (exploitation). Maximising the exploration term is equivalent to maximising the expected divergence between the expected posterior distribution, with and without observations expected under a policy - maximising this leads to behaviour that actively seeks out observations that resolve the most (posterior) uncertainty. Maximising the exploitation term is equivalent to changing policies to produce those observations that best match the agent’s prior beliefs about observations (a.k.a., its preferences), which is specified in the agent’s generative model. Hence, active inference balances exploration and exploitation, ensuring that an optimal agent pursues both. Often, in situations where the agent is uncertain about hidden states that are relevant to preferred observations, active inference agents will first perform more epistemically driven

actions to resolve uncertainty, before opting for a more pragmatic action that maximises utility, i.e., exploit the resolved structure of the environment.

Learning occurs in active inference by updating model parameters, such as the likelihood distributions and state transition beliefs. In the discrete state-space models commonly used in active inference, these likelihood and transition distributions are described as categorical distributions with matrices of parameters. These distributions are often equipped with conjugate Dirichlet priors [48], whose parameters take the form of “pseudocounts” or positive real numbers that parameterise prior beliefs about the corresponding categorical parameters. The values of these Dirichlet hyper parameters can be interpreted as “pseudocounts” that are proportional to the prior probability of seeing particular state-outcome contingencies or coincidences between states and actions over time. Learning is thus cast as posterior inference over these Dirichlet hyperparameters [48]. Hence, when a new observation is received by the agent, a posterior distribution over the model parameter is acquired to be used as the prior distribution in the next time step, equipping the agent to sequentially update beliefs about the model parameters. A learning rate parameter can also be specified to control how much the values in the Dirichlet distribution change after each time step, representing how quickly the agent can get stuck in its ways during learning [48].

To summarise, when there is a mismatch between the agent’s predictions and sensory inputs, the agent (i) updates its internal model to reduce future surprise by updating its beliefs about the states that caused the observation, and/or (ii) updating its beliefs about the dynamics of the world (updating model parameters), and/or (iii) actively engages with the environment to generate and maximise model evidence, thereby reducing future surprise. These processes of minimising surprise respectively map onto three fundamental processes: (i) perception, (ii) learning, and (iii) action.

3.2 Generative Model of Agential Learning

The Agential Learning Task

In this section, a discrete-time generative model of the classic free-operant agential learning task is presented as in Figure 1), along with a set of simulations presented in Figures 3 and 4). In the learning task, the agent produces an action (or not) by pressing a button (or not), and according to the objective contingency, an outcome is present or absent. In some conditions, the agent has 100% and 80% control over the outcome, referred to as the deterministic and probabilistic condition, respectively. In other conditions, the agent has no control, i.e., the outcome is produced at random, independent of the agent’s actions.

Generative Model of Agential Learning Task

The generative model an active inference agent is equipped with is in the form of a Partially Observable Markov Decision Process (POMDP; [45]). POMDPs express the generative model with a sequence of hidden states (s) that evolve over time. The hidden states inferred by the agent in this agential learning task are the objective contingency (positive, negative, or zero) between the self-produced

action and outcome, which is the context (or experimental condition) the agent is in ($s_t^{context}$).

At each time step (t), the current state is conditionally dependent on the state at the previous time step and on the actions (u ; aka control states) currently being executed. The actions are dependent on the policy (π ; aka action sequence) currently being executed. Each time step is associated with an observation (o) that depends only on the state at that time. The observations the agent receives are of the outcome ($o_t^{outcome}$), wherein the outcome can be present or absent, and the observations of the action the agent conducted (o_t^{action}), wherein the agent observes that it pressed the button or not.

The hidden and control states are classified into state factors, and observations are classified into observation modalities. This means that at any given time, observations will be evinced from each modality, and hidden states will be inferred from each state factor, and an action (control state) is selected accordingly. The s , u , and o are discrete random variables, so all model parameters are categorical distributions too.

The agent’s generative model is equipped with model parameters denoted as **A**, **B**, **C**, and **D** tensors that allow the agent to perform active inference. The likelihood tensor (**A**), represents the beliefs of probability of some observation given the states in the agent’s environment, $P(o_t|s_t)$. The top-left matrix in **A** tensor panel in Figure 2 illustrates that the agent believes the probability it will observe the outcome being present given it is in the positive control state and has pressed the button is 0.6. This value is not 1.0 as the agent cannot have already learned the precise likelihood mappings as it does not know the objective contingency in all of the possible contexts in which it could be operating, so it conducts learning by updating the likelihood tensor regarding outcomes via Dirichlet counts ($Dir(A^{outcome})$). The degree of control the agent perceives is indicated by the posterior probability of the state.

The state transition tensor (**B**), represents the beliefs of the dynamics of the environment as how hidden states and actions determine subsequent hidden states, $P(s_t|s_{t-1}, u_t)$. The objective contingency does not change over a block of trials, and we assume the agent knows this fact veridically, and thus their generative model has an identity matrix in the left matrix presented in the **B** tensor panel in Figure 2.

Sampling the environment occurs as a function of preferring each observation, represented in the preference tensor (**C**) in Figure 2, reducing uncertainty. There is a slight preference for not producing an action as producing an action costs resource. To introduce evidence variance, periods of sub-optimal action would be intentionally conducted by the agent to create variation in the observations and assess the agent’s generative model. The **D** tensor represents the agent’s beliefs of prior probability of being in each state, which is a flat distribution to reflect the agent’s lack of a bias towards being in a positive, negative, or zero-contingency state.

Simulation Results

All simulations described in this paper were conducted using the `sparse_likelihoods` 111

branch of `pymdp`, a freely available Python package for performing active inference in discrete state spaces [49]. The code used for the simulations described in this paper can be found here: https://github.com/riddhipits/iwai_agency_oneagent.

Figures 3 and 4 illustrate the results of simulations of an agent conducting agential learning in the deterministic and probabilistic learning task across three experimental conditions. Panels correspond to each experimental condition: positive control, negative control, and zero control. The three sub-panels in each panel illustrate the agent’s beliefs over time (x -axis) regarding the experimental condition (or context), the actions it took, and the outcomes it observed. The strength of the belief is reflected in the grayscale cells, with black cells indicating a value of 0.0 and white cells indicating a value of 1.0. The agent had 50 trials to learn the degree of control it had.

In the deterministic (100%) agency simulation (Figure 3), in the positive and negative control condition, the agent quickly learned that it had full positive and negative control, respectively; this is illustrated by the gradual transition from black to white on the top-most sub-panels. In the first few trials, the agent tracks (via Dirichlet counts) the outcome observation given the states it observes (actions) and infers (context) and reflects its learning of the environment being deterministic by updating the likelihood tensors in its generative model. Accordingly, the agent then infers, with certainty that it is in the positive or negative condition. As predicted, the agent introduces evidence variance by occasionally acting sub-optimally to increase certainty regarding its beliefs. In the probabilistic (80%) agency simulation (Figure 4), the agent learns similarly, albeit less quickly and with more uncertainty as illustrated with more grey cells.

In the zero-control condition, the agents in both simulations (Figure 3 and Figure 4) takes longer learn that its actions have no control over the outcome. To elaborate, in Figure 3, near the first few trials (Box A), the agent’s actions of pressing the button were coincidentally paired with the outcome being present, which is why it had a higher belief of being in the positive control state. And in the middle of the block of trials (Box B), the agents actions aligned with what it would predict to perceive in a negative control condition, which is why its beliefs shift towards the negative control condition until it receives evidence against that belief. Finally, the agent’s beliefs increase for the zero-control condition. Throughout the block of trials, the agent tests its hypotheses by variably pressing the button or not.

4 Discussion and Concluding Remarks

These simulations reveal that the active inference framework has great potential to provide a comprehensive model of agential learning as it intricately ties perception, actions, and learning processes involved in learning one’s degree of agency, describing the interactions between them resulting from the minimisation of a single metric: free energy. Previous models have treated these processes as optimising disparate functions.

Compared to Bayesian agents, active inference agents possess a deeper representation of the causal structure and dynamics of the environment as an active inference agent’s generative model is equipped with beliefs about state transitions across time. This is leveraged by the active inference agent as it allows the agent to consider future states and observations based on future actions to optimally select an action. The actions maximise evidence for the agent’s generative model of their environment by exploring the environment when uncertainty is high and then exploiting the environment to attain preferred observations/outcomes, and introduce evidence variance to continually assess the agent’s generative model.

The active inference model of agential learning may allow us to explain individual differences in agential learning. For example, agents experiencing learned helplessness (a key symptom of depression) may have a higher learning rate for the zero-control state due to generalisation from trauma, resulting in them having a bias and getting stuck when the belief of being in a zero-control state is higher. Over time, this may result in them developing a habit of not producing an action (due to deep temporal active inference models; see [50]), resulting in reduced variance in sampling the environment.

The simulation results in this paper emphasise that observations of different action-outcome combinations make a big difference to the perceived contingency in a zero-control condition. This predicts that agents who produce actions would experience more (but accidental) action-outcome-present observations and thereby perceive an illusion of (positive) control, whereas agents who withhold actions would perceive more no-action-outcome-present observations, resulting in perceiving zero control. The predictions are in line with data from humans as experimenters showed that non-depressed individuals produced more actions in the zero-control condition, perceiving an illusion of (positive) control, and individuals experiencing depression withheld actions, perceiving a lack a control, potentially explaining their lack of sense of agency [51].

Nonetheless, further examination of the active inference formulation of agential learning is warranted. In future research studies, we intend to: (i) conduct statistical model comparisons between the different accounts of agential learning via model fitting to human behavioural data, (ii) examine if active inference explains individual differences in agential learning across the depression spectrum, and (iii) explore more complex scenarios of agential learning such as one with multiple agents and outcomes.

5 Figures

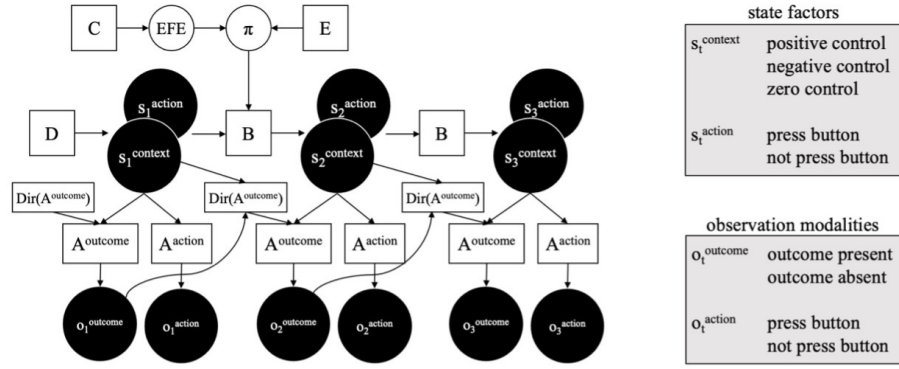


Fig. 1. A graphical representation [52] of the active inference based generative model of the agential learning task. The variables of the model are illustrated as circles and model parameters as squares and rectangles. The arrows indicate the direction of influence. Please see the main text for a description of the variables and parameters.

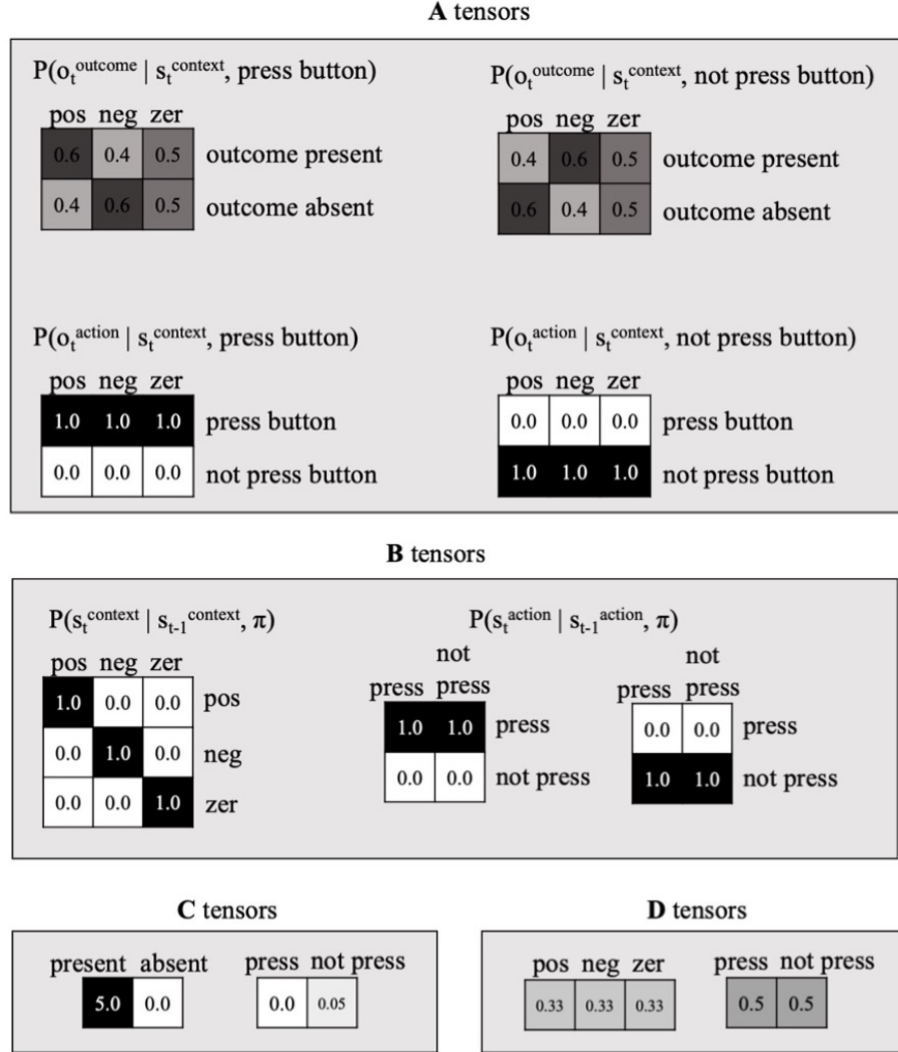


Fig. 2. The details of the model parameters of the generative model of the agential learning task.

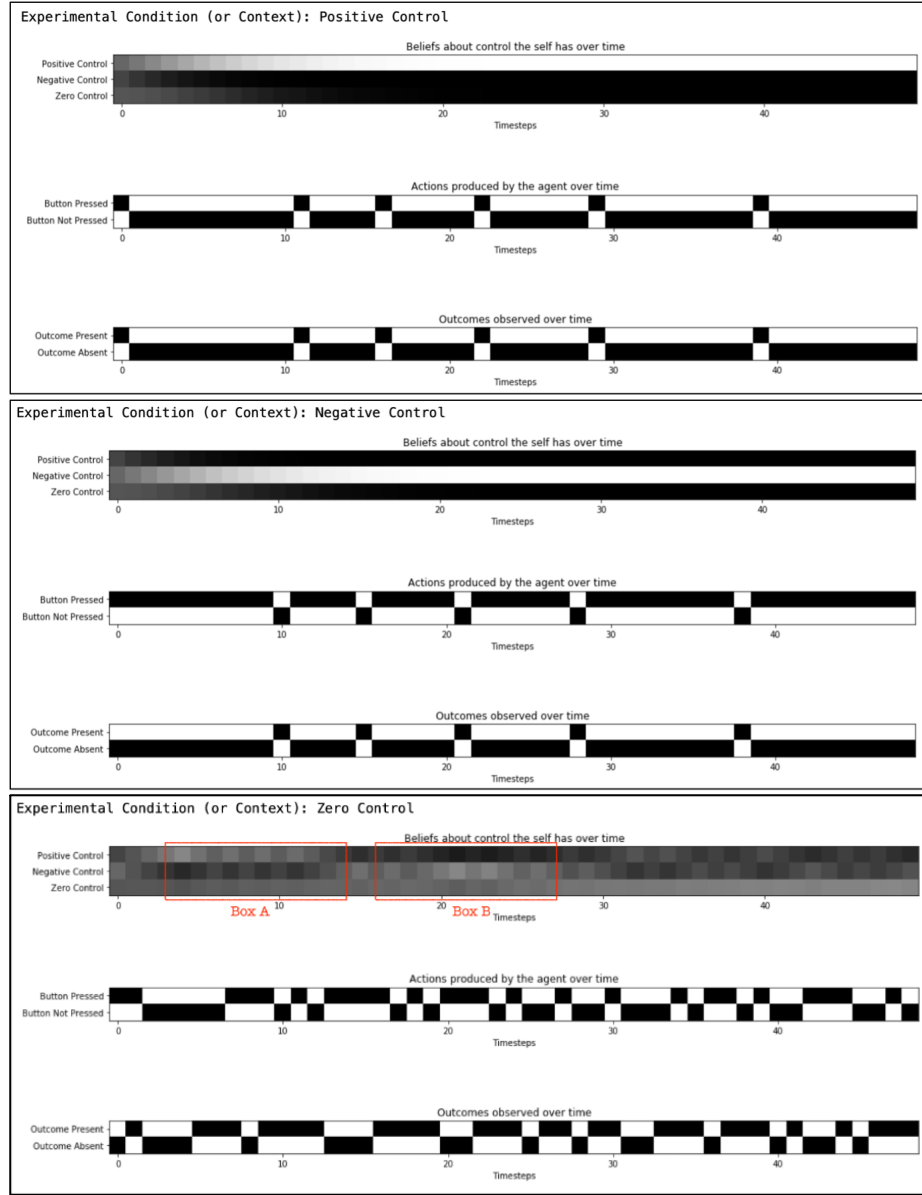


Fig. 3. Simulation results for deterministic (100% control) agential learning task. The three panels illustrate three separate simulations, one for each experimental condition: positive control, negative control, and zero control. Within each panel of simulation result, there are three sub-panels, where the x axis is the timestep. The black cells represent the value of 0.0 and white cells represent the value of 1.0, so the grayscale cells are values within that range. The top sub-panel illustrates the beliefs the agent has regarding the context states, the middle sub-panel illustrates the actions the agent selected over time, and the bottom sub-panel illustrates the outcomes the agent observed over time.

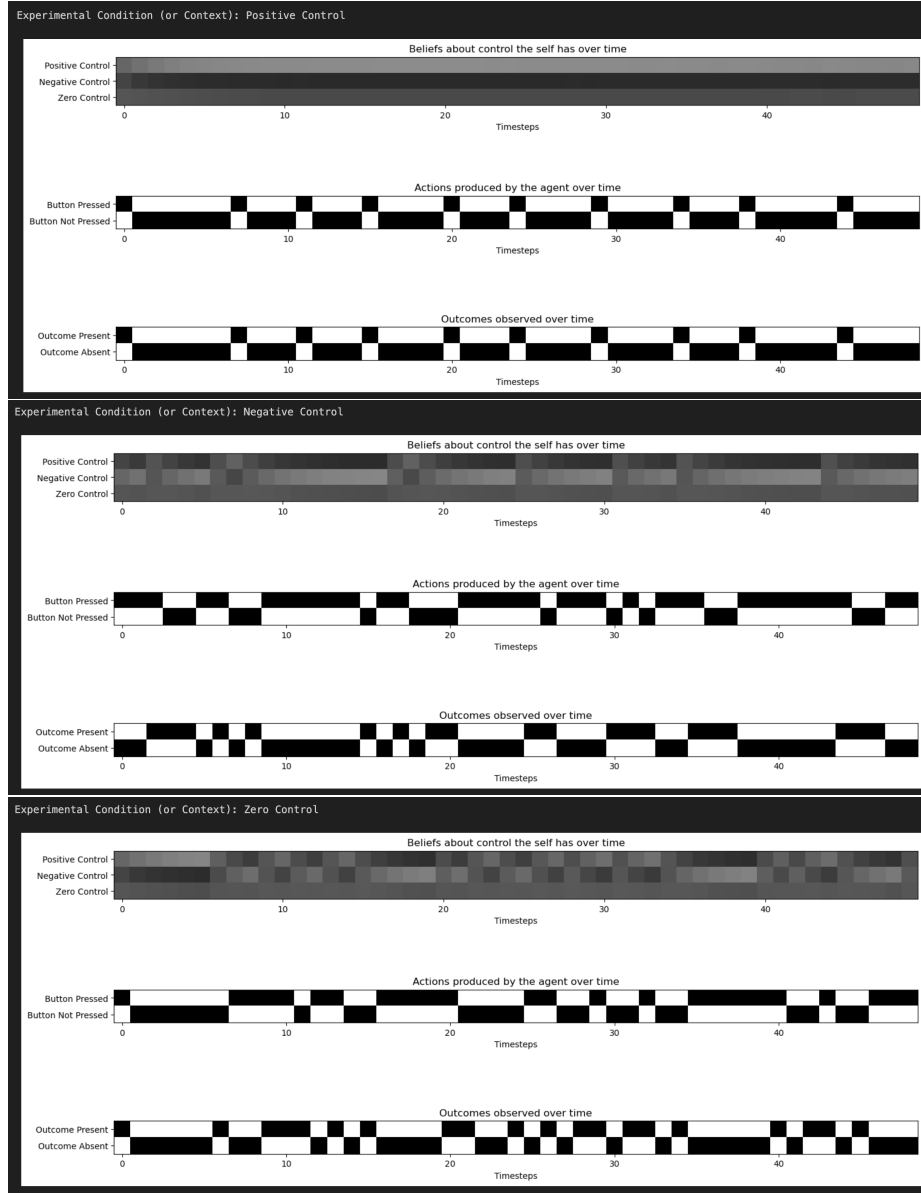


Fig. 4. Simulation results for probabilistic (80% control) agential learning task. The three panels illustrate three separate simulations, one for each experimental condition: positive control, negative control, and zero control. Within each panel of simulation result, there are three sub-panels, where the x axis is the timestep. The black cells represent the value of 0.0 and white cells represent the value of 1.0, so the grayscale cells are values within that range. The top sub-panel illustrates the beliefs the agent has regarding the context states, the middle sub-panel illustrates the actions the agent selected over time, and the bottom sub-panel illustrates the outcomes the agent observed over time.

References

- [1] S. Gallagher, “Philosophical conceptions of the self: Implications for cognitive science,” *Trends in cognitive sciences*, vol. 4, no. 1, pp. 14–21, 2000.
- [2] P. Haggard, “Sense of agency in the human brain,” *Nature Reviews Neuroscience*, vol. 18, no. 4, pp. 196–207, 2017.
- [3] M. Albarracin and R. J. Pitliya, “The nature of beliefs and believing,” *Frontiers in Psychology*, vol. 13, 2022.
- [4] P. F. Verschure, C. M. Pennartz, and G. Pezzulo, “The why, what, where, when and how of goal-directed choice: Neuronal and computational principles,” *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1655, p. 20130483, 2014.
- [5] C. B. Ferster, “The use of the free operant in the analysis of behavior,” *Psychological Bulletin*, vol. 50, no. 4, p. 263, 1953.
- [6] L. G. Allan and H. M. Jenkins, “The judgment of contingency and the nature of the response alternatives,” *Canadian Journal of Psychology/Revue canadienne de psychologie*, vol. 34, no. 1, p. 1, 1980.
- [7] D. R. Shanks and A. Dickinson, “Instrumental judgment and performance under variations in action–outcome contingency and contiguity,” *Memory & Cognition*, vol. 19, pp. 353–360, 1991.
- [8] E. A. Wasserman, D. Chatlosh, and D. Neunaber, “Perception of causal relations in humans: Factors affecting judgments of response–outcome contingencies under free–operant procedures,” *Learning and motivation*, vol. 14, no. 4, pp. 406–432, 1983.
- [9] E. A. Wasserman, S. M. Elek, D. L. Chatlosh, and A. G. Baker, “Rating causal relations: Role of probability in judgments of response–outcome contingency,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 19, no. 1, p. 174, 1993.
- [10] F. Vallée-Tourangeau, R. A. Murphy, and A. Baker, “Contiguity and the outcome density bias in action–outcome contingency judgements,” *The Quarterly Journal of Experimental Psychology Section B*, vol. 58, no. 2b, pp. 177–192, 2005.
- [11] F. Vallee-Tourangeau and R. Murphy, “Action-effect contingency judgment tasks foster normative causal reasoning,” in *Proceedings of the Twenty First Annual Conference of the Cognitive Science Society*, 1999, pp. 820–820.
- [12] R. M. Msetfi, R. A. Murphy, J. Simpson, and D. E. Kornbrot, “Depressive realism and outcome density bias in contingency judgments: The effect of the context and intertrial interval,” *Journal of Experimental Psychology: General*, vol. 134, no. 1, p. 10, 2005.
- [13] P. W. Cheng, “From covariation to causation: A causal power theory,” *Psychological review*, vol. 104, no. 2, p. 367, 1997.
- [14] D. Hume, “A treatise of human nature: Volume 1: Texts,” 1739.
- [15] I. Kant, “Critique of pure reason. 1781,” *Modern Classical Philosophers*, Cambridge, MA: Houghton Mifflin, pp. 370–456, 1908.
- [16] A. Michotte, *The perception of causality*. Routledge, 2017, vol. 21.

- [17] D. R. Shanks, F. J. Lopez, R. J. Darby, and A. Dickinson, “Distinguishing associative and probabilistic contrast theories of human contingency judgment,” in *Psychology of learning and motivation*, vol. 34, Elsevier, 1996, pp. 265–311.
- [18] J. De Houwer and T. Beckers, “A review of recent developments in research and theories on human contingency learning,” *The Quarterly Journal of Experimental Psychology: Section B*, vol. 55, no. 4, pp. 289–310, 2002.
- [19] O. Pineño and R. R. Miller, “Comparing associative, statistical, and inferential reasoning accounts of human contingency learning,” *Quarterly Journal of Experimental Psychology*, vol. 60, no. 3, pp. 310–329, 2007.
- [20] D. R. Shanks, “Associationism and cognition: Human contingency learning at 25,” *Quarterly Journal of Experimental Psychology*, vol. 60, no. 3, pp. 291–309, 2007.
- [21] N. J. Mackintosh, “A theory of attention: Variations in the associability of stimuli with reinforcement.,” *Psychological review*, vol. 82, no. 4, p. 276, 1975.
- [22] R. R. Miller and L. D. Matzel, “The comparator hypothesis: A response rule for the expression of associations,” in *Psychology of learning and motivation*, vol. 22, Elsevier, 1988, pp. 51–92.
- [23] J. M. Pearce and G. Hall, “A model for pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli.,” *Psychological review*, vol. 87, no. 6, p. 532, 1980.
- [24] R. A. Rescorla, “A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement,” *Classical conditioning, Current research and theory*, vol. 2, pp. 64–69, 1972.
- [25] A. R. Wagner and R. A. Rescorla, “Inhibition in pavlovian conditioning: Application of a theory,” *Inhibition and learning*, pp. 301–336, 1972.
- [26] G. B. Chapman, “Trial order affects cue interaction in contingency judgment.,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 17, no. 5, p. 837, 1991.
- [27] J. De Houwer and T. Beckers, “Higher-order retrospective revaluation in human causal learning,” *The Quarterly Journal of Experimental Psychology Section B*, vol. 55, no. 2b, pp. 137–151, 2002.
- [28] A. Dickinson, “Within compound associations mediate the retrospective revaluation of causality judgements,” *The Quarterly Journal of Experimental Psychology: Section B*, vol. 49, no. 1, pp. 60–80, 1996.
- [29] P. W. Cheng and L. R. Novick, “Covariation in natural causal induction.,” *Psychological review*, vol. 99, no. 2, p. 365, 1992.
- [30] F. J. López, J. Almaraz, P. Fernández, and D. Shanks, “Adquisición progresiva del conocimiento sobre relaciones predictivas: Curvas de aprendizaje en juicios de contingencia,” *Psicothema*, pp. 337–349, 1999.
- [31] K. Friston, T. FitzGerald, F. Rigoli, P. Schwartenbeck, and G. Pezzulo, “Active inference: A process theory,” *Neural computation*, vol. 29, no. 1, pp. 1–49, 2017.

- [32] F. Blanco, H. Matute, and M. A. Vadillo, “Mediating role of activity level in the depressive realism effect,” 2012.
- [33] F. Blanco, H. Matute, and M. A. Vadillo, “Interactive effects of the probability of the cue and the probability of the outcome on the overestimation of null contingency,” *Learning & Behavior*, vol. 41, pp. 333–340, 2013.
- [34] N. Byrom, R. Msetfi, and R. Murphy, “Two pathways to causal control: Use and availability of information in the environment in people with and without signs of depression,” *Acta psychologica*, vol. 157, pp. 1–12, 2015.
- [35] T. L. Griffiths and J. B. Tenenbaum, “Structure and strength in causal induction,” *Cognitive psychology*, vol. 51, no. 4, pp. 334–384, 2005.
- [36] M. R. Waldmann, “Competition among causes but not effects in predictive and diagnostic learning,” *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 26, no. 1, p. 53, 2000.
- [37] J. K. Kruschke, “Bayesian approaches to associative learning: From passive to active learning,” *Learning & behavior*, vol. 36, no. 3, pp. 210–226, 2008.
- [38] J. B. Tenenbaum, T. L. Griffiths, and C. Kemp, “Theory-based bayesian models of inductive learning and reasoning,” *Trends in cognitive sciences*, vol. 10, no. 7, pp. 309–318, 2006.
- [39] N. Chater, M. Oaksford, U. Hahn, and E. Heit, “Bayesian models of cognition,” *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 1, no. 6, pp. 811–823, 2010.
- [40] K. Doya, S. Ishii, A. Pouget, and R. P. Rao, *Bayesian brain: Probabilistic approaches to neural coding*. MIT press, 2007.
- [41] H. Von Helmholtz, *Handbuch der physiologischen Optik*. Voss, 1867, vol. 9.
- [42] J. D. Nelson, “Finding useful questions: On bayesian diagnosticity, probability, impact, and information gain,” *Psychological review*, vol. 112, no. 4, p. 979, 2005.
- [43] G. De Ath, R. M. Everson, A. A. Rahat, and J. E. Fieldsend, “Greed is good: Exploration and exploitation trade-offs in bayesian optimisation,” *ACM Transactions on Evolutionary Learning and Optimization*, vol. 1, no. 1, pp. 1–22, 2021.
- [44] K. Friston, “The free-energy principle: A unified brain theory?” *Nature reviews neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.
- [45] T. Parr, G. Pezzulo, and K. J. Friston, *Active inference: the free energy principle in mind, brain, and behavior*. MIT Press, 2022.
- [46] K. J. Friston, J. Daunizeau, and S. J. Kiebel, “Reinforcement learning or active inference?” *PloS one*, vol. 4, no. 7, e6421, 2009.
- [47] K. Friston, F. Rigoli, D. Ognibene, C. Mathys, T. Fitzgerald, and G. Pezzulo, “Active inference and epistemic value,” *Cognitive neuroscience*, vol. 6, no. 4, pp. 187–214, 2015.
- [48] R. Smith, K. J. Friston, and C. J. Whyte, “A step-by-step tutorial on active inference and its application to empirical data,” *Journal of mathematical psychology*, vol. 107, p. 102632, 2022.

- [49] C. Heins, B. Millidge, D. Demekas, *et al.*, “Pymdp: A python library for active inference in discrete state spaces,” *arXiv preprint arXiv:2201.03904*, 2022.
- [50] K. J. Friston, R. Rosch, T. Parr, C. Price, and H. Bowman, “Deep temporal models and active inference,” *Neuroscience & Biobehavioral Reviews*, vol. 90, pp. 486–501, 2018.
- [51] F. Blanco, H. Matute, and M. A. Vadillo, “Depressive realism: Wiser or quieter?” *The Psychological Record*, vol. 59, no. 4, pp. 551–562, 2009.
- [52] K. J. Friston, T. Parr, and B. de Vries, “The graphical brain: Belief propagation and active inference,” *Network neuroscience*, vol. 1, no. 4, pp. 381–414, 2017.