Bachelor thesis

Artificial Intelligence

Radboud University

Simulating Affect-Biased Attention in Active Inference

Author:
Jort Gutter
s1017170

First supervisor:
E.J. de Wolff, MSc.
Donders Institute for Brain, Cognition and Behavior
erwin.dewolff@donders.ru.nl

Second supervisor:
dr. J.H.P. Kwisthout
Donders Institute for Brain, Cognition and Behavior
j.kwisthout@donders.ru.nl

June 25, 2021

Abstract

Following a recent challenge to the active inference account, we investigate in this paper whether it can reproduce affect-biased attention. To do so, we built a single-layer predictive processing agent. By introducing a highly negative, but unlikely stimulus at some point during the simulation, we measured whether and how the agent would alter its behavior in order to avoid that unpleasant event in the future. Additionally, we explored how this change depended on the level of sensory modalities, the amount of experience, or the salience of the encounter. Our results show evidence for such long-term changes in behavior, which opens up the road for future research to extend our findings to more complete models of attention.

Contents

Chapter 1

Introduction

One of the aims of artificial intelligence is to understand the underlying processes of the human brain. The thought is that if we understand how the brain manages to solve complex problems, we might gain insights in how to build machines to do the same. Vice versa, insights from artificial intelligence are used by cognitive neuroscience to help develop theories of the inner workings of the human brain [11]. In this context the brain is viewed as an information processing machine, or computer [15], which processes incoming information through the senses and generates output in the form of behavior, or action. Over the years, various theories have been put forward in an attempt to explain how the brain accomplishes this.

One such unifying theory which attempts to explain all cognition processes and has become popular in recent years is predictive processing (PP) [8]. PP posits that the purpose of the brain is to minimize surprise by predicting and acting upon the world.

The active inference (AI) paradigm, which extends the PP by including action, attempts to explain why and how agents act based on prior expectations and predictions about the world. It unifies prediction and action in a single concept, where updating your predictions is aligning your model with the world, and acting is aligning the world with your model [1].

An active inference agent tries to reach it goals (e.g. not being hungry) by predicting that it has reached that goal (it has a full stomach). It will then perform the action that minimizes the difference between its predicted (future) perceptions (the agent's sensation of hunger) and its goals (the agent is no longer hungry). These predictions are generated by the agent's internal generative model of the world [3].

According to the theory, notions of reward or utility which are often used in more classical approaches of learning, like reinforcement learning, are not required. Instead, it is claimed that the notion of surprise minimization can fully account for all cognitive functions and behavior.

PP has found many applications, both in neuroscience [9, 2] and artificial intelligence [5].

However, in a recent paper (2020), Ransom et al. argue that PP can not account for all cognitive processes, and in particular fails to explain the notion of affect-biased attention [13]. This type of attention is thought to play a role in emotion regulation [16]. Ransom et al. explain affect-biased attention with the example of a dog:

Suppose you walk by a certain yard every day on your way to work. One day, there is a large dog in that yard, barking at you as you pass by. This startles you, and on subsequent walks by that yard, your attention is drawn to that particular fence, even though you know that the presence of that dog is a rare event, as it only happened once [13].

To summarize, affect-biased attention is thus attention to aspects of the environment that generated a rare, but salient observation in the past. Affect-biased is distinguished from endogenous attention, which is attention by voluntary control, as the attention to the yard happens involuntary. It is also distinguished from exogenous attention, as it is not exclusively generated by current sensory input [12].

According to Ransom et al., the reason that PP can not account for this type of attention, is that PP tries to maximise precision instead of optimizing utility with utility functions, while salience and precision are not always aligned. Even in cases where the expected precision of prediction errors can be high,like in the case of the dog, the reward or punishment associated with the unlikely event might still be high enough to warrant attention [14]. Therefore, they claim, it can not be the case that prediction minimization with PP can fully explain all cognitive processes and phenomena [14, 13].

This criticism raised by Ransom et al. is a severe limitation for PP, as it claims to be an all encompassing theory of cognition [3, 4, 6], which means it should be able to account for phenomena such as affect-biased attention. This means that we either have to show that PP can actually account for affect-biased attention, which would refute the claim made by Ransom et al., or we will have to extend the PP framework to overcome this limitation. If both approaches fail, PP will have to be rejected as a unified theory of the mind.

With this study we aimed to investigate methods to account for affect-biased attention within the active inference framework by addressing the following question:

> *How can active inference account for affect-biased attention in a simulated food-collection task?*

To answer this question we programmed an active inference agent which can act in a small environment. We let the agent learn the model parameters for a number of iterations, after which we presented the agent with a single salient and undesirable observation. We investigated whether there was any significant or long-term change of behavior by comparing the actions of the agent before and after this event. We hypothesized that if the salience of an observation is high enough, the agent could be affected long term. By varying the salience, timing and availability in the number of sensory modalities of this event, we explored how these parameters are linked to the extend of behavioral change.

By demonstrating such a long term behavioral change we could make a first step towards a working model of affect-biased attention that fits within the active inference, or predictive processing framework. In the following chapters, we will describe how an active inference agent learns and acts (preliminaries), and how we set up the environment and the agent, which parameters we manipulated and how that affected the agent's behavior (research). Finally, we will discuss the implications of these results on our research question, and go over the limitations of our methods (conclusions).

Chapter 2

Preliminaries

2.1 Active inference

In this chapter we will look more closely at how the generative model allows active inference agents to learn and act in an environment, and how the environment generates observations for the agent through the generative process.

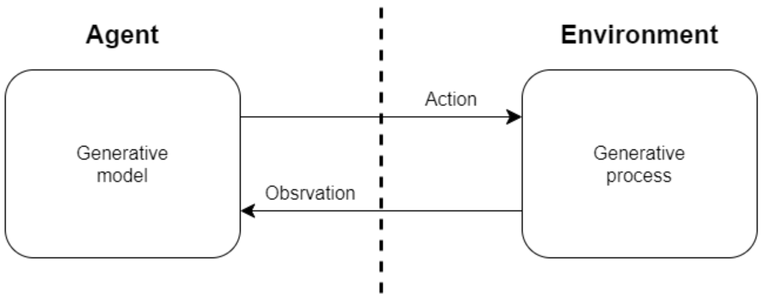

Figure 2.1: The interaction between an agent and the environment, through the generative model and the generative process

2.1.1 The generative process

When an agent performs an action, the world will end up in a certain state. This state generates an observation which the agent perceives (See figure 2.1). The generative process is how the environment generates these observations. There is not a set way in which observations are generated, and this process can be extremely complex. Agents don't have direct access to the true hidden states of the environment. This means that the agent generally does not know in which state it is at any given moment, but has to infer that from the observations.

2.1.2 The generative model

The generative model is the internal model of an agent of the external generative process. It generates predictions over future observations based on learned model parameters. Active inference agents generate these predictions and actions by minimizing *free energy* or *surprise* of observations of which free energy is an upper bound [10, 7]. This can be done in two ways:

1. By updating the internal beliefs of the agent. This is done by updating the model parameters to align them to observations of the generative process. This is called perceptual inference.

2. By performing an action which will align the external world with the internal expectations. In this case, the agent minimizes variational free energy, or the expected surprise over future observations.

Another way to express the minimization of variational free energy is the maximization of the quality of a policy π, $\mathbf{Q}_\tau(\pi)$:

$$\underbrace{\mathbf{Q}_\tau(\pi)}_{\text{Quality}} = \underbrace{E_{Q(o_\tau|\pi)}\left[\ln P\left(o_\tau \mid m\right)\right]}_{\text{Extrinsic value}} + \underbrace{E_{Q(o_\tau|\pi)}\left[D\left[Q\left(s_\tau \mid o_\tau, \pi\right) \| Q\left(s_\tau \mid \pi\right)\right]\right]}_{\text{Epistemic value}}$$

We see that the quality can be expressed by the sum of two terms:

1. **Extrinsic value**:
   This term represents the utility of the expected outcome given a policy. It is calculated by multiplying the predicted outcome of an action $E_{Q(o_\tau|\pi)}$ by the log of the expected outcome $P(o_\tau \mid m)$. The expected outcome expresses the goals of the agent. Therefore, when the extrinsic value is high, the agent is more likely to favor exploitation over exploration.

2. **Epistemic value**:
   When this second term dominates the equation, the agent favors exploration, as it represents the expected information gain. It is calculated by multiplying the predicted outcome by the entropy between the posterior distribution of possible observations, and the distribution over predicted states. When the agent has absolute certainty as to which state it will end up in given a policy, this term drops to zero.

From the resulting quality distribution an action is sampled. The interaction of these two terms will cause an agent that is uncertain about its environment to first explore, until the epistemic value drops. With the extrinsic value now dominating the equation, the agent will start to select actions based on its internal goals.

Chapter 3

Research

3.1 General setup

In order to investigate if and how a single, salient experience can alter the behavior of an active inference agent, we created the following setup:

We coded a simple active inference agent in python (For the code, see Appendix A). At each point in time, the agent could choose one out of three actions, which depended on its model of the world and its preferred outcomes. Each action would put the agent in one out of three states in the environment. Thereafter, the generative process generated categorical outcomes over three separate sensory modalities given the chosen sate. These were then given back to the agent, which it used to update its internal beliefs about the environment. We will elaborate more on the exact details of how the agent and the environment were set up in the following sections.

At some point in time, after the agent has sufficiently learned the model parameters, we presented it with a single, salient, and undesirable experience. Through further iterations we observed whether there was any behavioral change after this point in time.

3.1.1 Environment

The environment, or the generative process, is what the generative model of the agent tries to learn. As we have explained above, the agent performs an action, which gets translated to a state by the environment. The generative process will then generate an observation, which it will then send back to the agent.

For our experiments we purposely kept the generative process simple. We created three states, which were linked one-to-one to the three actions the agent could perform. This means that the selection of states by the agent was deterministic, and the agent did not have any uncertainty about which state it would end up in.

For each state, the generated observation consisted of three sensory modalities. For each modality the generative process picked a value out of three categorical outcomes, based on probability weights.
We chose these probability weights to be the same over all sensory modalities for simplicity. These outcomes were drawn using preset probabilities of each outcome given the sensory modality and state, which are shown in table 3.1. The weights of outcome two are set to zero, as this outcome represents the salient encounter, and we manually decide per experiment when these outcomes are picked.

state	outcome 0	outcome 1	outcome 2
s0	0.6	0.4	0
s1	0.9	0.1	0
s2	0.3	0.7	0

Table 3.1: Probability weights for each observation outcome of a sensory modality given the state

3.1.2 Agent

As we have explained earlier on, the generative model of active inference agents can choose actions by calculating the quality over all policies based on internal states, which it will then use to sample an action. For our experiment we gave the agent the following internal parameters:

1. **Expected outcome**
   This 3 by 3 matrix (see table 3.2) represented the internal goals by the agent.The expected outcomes summed to 1 and were the same for each sensory modality for simplicity of the model.

   The first outcome for each modality was the preferred outcome by the agent. The second outcome was somewhat undesirable and the third was highly undesirable, representing the salient unpleasant encounter.

   In other words, the agent preferred a distribution over the observations where the first outcome was the most likely, and the third almost never occurred.

2. **Predicted outcome**
   This was a 3 by 3 by 3 matrix. It held the predicted distributions over the observations, given the state and sensory modality. These values were kept as counts; for each observation, the count of a given observation was increased by the value of a corresponding salience

matrix. When calculating the quality, these counts were transformed into a probability distribution for each combination of a state and sensory modality.

3. **Salience matrix**
   This small vector of size 3 represented the salience of each possible outcome, and influenced the rate of learning of the predicted outcomes.

   The first two outcomes typically had a salience value of 1, with the third varying between experiments.

4. **modality weights**
   This vector indicated the weight each sensory modality had in deciding the final quality distribution. These weights varied between experiments, but always summed to 1.

Sensory modality	outcome 0	outcome 1	outcome 2
modality 0	0.9	0.0999	0.00001
modality 1	0.9	0.0999	0.00001
modality 2	0.9	0.0999	0.00001

Table 3.2: Expected outcome per sensory modality

Using these parameters, the agent calculated a quality vector for each sensory modality as explained in the preliminaries chapter. See Appendix A for the precise implementation. By applying softmax to these vectors we ensured we are dealing with a proper probability distributions. By weighing the results by the modality weights, we created a final distribution of qualities over the policies. From these, an action was drawn, which generated an observation. Finally, the observation was weighed by the salience matrix and used to update the predicted outcome, or internal beliefs about the world.

This cycle was repeated for as many iterations as the experiments required, logging the final quality distribution at each point in time for later analysis.

3.2 Experiments

Over a number of iterations we let the agent choose and action and observe the world. During this time, the agent would learn the probabilities of each outcome in each state, and we expected it to start to favor one of its actions over the other ones. Our expectation was that this would be action 1, as the distribution of the outcomes over this state more is closely aligned to the expected outcomes, or desires, of the agent than the other two states (See table 3.1 and 3.2).

After a set number of these learning iterations, we presented a salient, unexpected and undesirable observation to the agent in its preferred state. In the experiments we referred to his encounter as *dog-day*, as a reminder of the example of the dog in the yard presented by Ransom et al. [13]. Dog-day always happened at the first moment in time after the set burn-in period that the agent performed action 1, which would land the agent in its most preferred state, state 1.

By varying several parameters we studied what the effect of these parameters was on the changed behavior of the agent. Over the course of three types of experiment we altered the following three parameters:

1. **Salience of the encounter**
   By changing the salience of the unpleasant observational outcome compared to the other two outcomes, we tested whether modelling salience like this would be sufficient in causing long-term behavior changes.

2. **Timing of the encounter** Here we varied the timing of the encounter, i.e. the number of action-perception cycles the agent makes before introduction of the salient event. This could show us the effect of familiarity with the environment on potential behavioral changes.

3. **The number of affected sensory modalities** This last parameter determined how many out of the three sensory modalities would be affected by the encounter. Since agents often get input in various ways in complex environments, it is important to investigate how the availability to the senses of such an encounter impact the effect.

3.2.1 Varying salience

For this first experiment we varied the salience, or the weight, of the unexpected encounter, compared to normal outcomes. For outcome 0 and outcome 1 we always set the salience to 1. We ran this experiment twenty times, each time with a different salience for outcome 2, the outcome presented to the agent on dog-day. We increased this salience linearly in steps of 50 from 0 (control) to 1000. All other variables were kept constant; dog-day happened after 10000 regular cycles, and only a single sensory modality was taken into account when calculating the policies.

For each of the different experimental conditions, we ran the simulation twenty times to account for probabilistic variation. The behavior of the agent in the 200 cycles leading up to dog-day and directly after dog-day were recorded.

3.2.2 Varying dog-day

Here we kept the salience of the dog encounter at a fixed value of 500. We let the agent run for 20000 cycles, while varying dog-day in steps of 2000 cycles from 0 up to 18000. The agent kept track of the qualities of its policies over time. We ran each of these ten experimental condition ten times, again to counter variation introduced by drawing from probability distributions.

3.2.3 Varying number of sensory modalities

In this final experiment, we varied the number of affected sensory modalities from none to all three. This means that when the dog encounter happens, the observation generated by the generative process was overwritten by outcome 2 for all affected sensory modalities, while the outcomes of the unaffected modalities were left unchanged.

The agent ran through 50000 cycles, with dog-day set to happen after cycle 15000. The salience of a dog encounter was kept at 500 over all runs, and per condition we ran the simulation ten times.

3.3 Results

3.3.1 Effects of varying salience

In order to investigate the potential effects of the salience of the dog encounter on the behavior of the agent, we looked at the fraction of performed actions that go to state 1, the most desirable state for the agent. By plotting the fraction over the 200 cycles before and directly after the dog encounter, we can visualize the immediate effect on the behavior. On top of all trials, we also plotted the mean over each condition. As shown in figure 3.1, There is a strong correlation between a decrease in the selection of action 1 and the salience of the encounter.

3.3.2 Effects of varying dog-day

In this experiment we kept track of the policy qualities over the entire run of a trial. Per action, we plotted the qualities of all runs. In addition, we included the mean of each experimental condition. When plotted (figure 3.2), we see that at all conditions there is an immediate change in policies after the dog encounter; the quality of the affected action, action 1, is decreased significantly in all cases. Action 0, the second best option for the agent, takes over most of the lost quality. We do see that the effects decrease logarithmically as the day of the dog encounter increases.

Additionally, the results show that the effects take a long time to wear off after the encounter. None of the qualities return to their original values within the run time of the experiment. Furthermore, we see that there is little variation between trials that shared the same initial parameters.

3.3.3 Effects of varying number of sensory modalities

In this final experiment we collected and plotted the results in the same way as in the previous experiment; we plotted the qualities of the actions over time, with different colors for different initial parameters. What is most striking about the result plotted in figure 3.3, is that there seems to be a linear relationship between the shift in quality and the number of affected sensory modalities. E.g., when two out of three modalities observe outcome 2, we see $\frac{2}{3}$ of the change in quality when comparing these results to the case where all three modalities observed outcome 2.

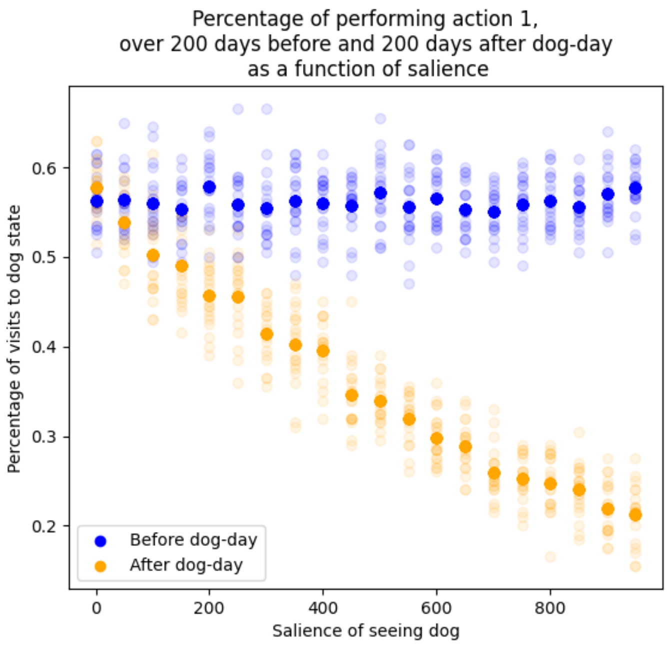

Figure 3.1: Results of varying the salience of a dog encounter on the selection of the affected action. The mean over each set of 10 identical initial conditions is represented by the fully saturated points. See Appendix A.2 for an enlarged version.

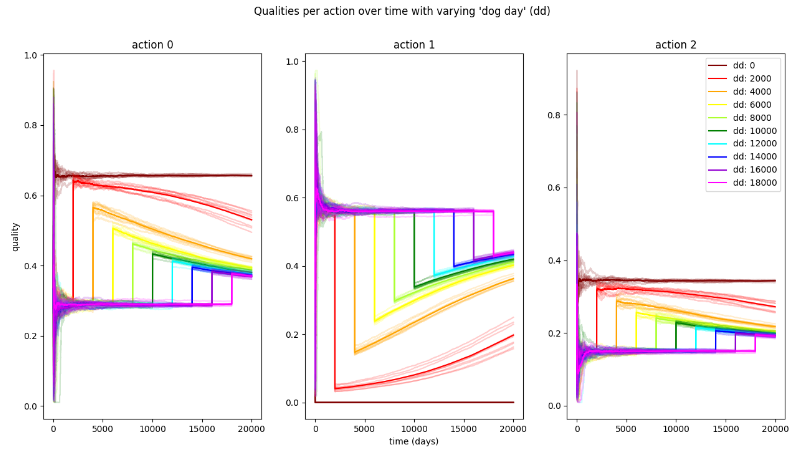

Figure 3.2: The qualities over time of all trials with varying dog days, for all three actions. The mean over each set of 10 similar trials is represented by a more saturated line. See Appendix A.2 for an enlarged version.

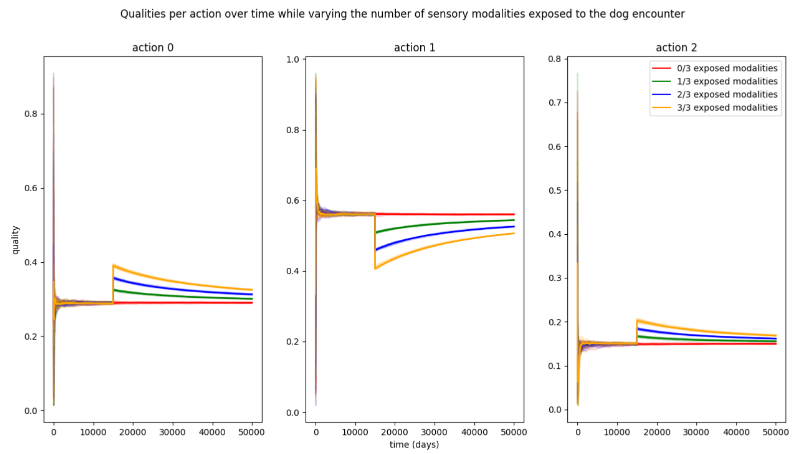

Figure 3.3: The qualities over time of all trials with varying amount of affected sensory modalities, for all three actions. The mean over each set of 10 similar trials is represented by a more saturated line. See Appendix A.2 for an enlarged version.

Chapter 4

Conclusions

Our work has shown that when a one-off event is salient enough, it can have long term effects on the behavior of agents within the active inference framework. Although our experimental setup was not an accurate model of attention, it suggested that there are ways for affect-biased stimuli to have a lasting impact within the framework on actions. If such effects can steer action, it is conceivable they could also steer attention.

If successful, such further research would refute the claim by Ransom et al. [13] that perceptive processing can not accounting for affect-biased attention. This would be further support for PP as an all-encompassing theory of cognition.

An underlying theory which is so grounded in mathematics as PP, could help us understand the principles behind intelligence. This would give us tools to apply these concepts in applications of artificial intelligence.

However, there are a number of limitations to our research. For one, our model only consists of a single layer. It has yet to be seen if similar effects can be achieved with more complex models. To properly model attention, a deep hierarchical model is necessary, and thus it is important that these effects can be generalized to such cases.

Furthermore, the actions of our agent were deterministic There was no uncertainty in which state the agent would actually end up after performing a given action. Because of this, the epistemic part of our quality had no influence on the policy selection. A next step would be to extend our experiments to such non deterministic environments.

Another direction of future research could be to tie the salience of observations to some inner state of the agent, be it the expected outcomes, or inner goals, or the prediction error.

Bibliography

[1] Rick A Adams, Stewart Shipp, and Karl J Friston, *Predictions not commands: active inference in the motor system*, Brain Structure and Function **218** (2013), no. 3, 611–643.

[2] Megan A Boudewyn, Debra L Long, and Tamara Y Swaab, *Graded expectations: Predictive processing and the adjustment of expectations during spoken language comprehension*, Cognitive, Affective, & Behavioral Neuroscience **15** (2015), no. 3, 607–624.

[3] Andy Clark, *Whatever next? predictive brains, situated agents, and the future of cognitive science*, Behavioral and brain sciences **36** (2013), no. 3, 181–204.

[4] ———, *Surfing uncertainty: Prediction, action, and the embodied mind*, Oxford University Press, 2015.

[5] Johan Engström, Jonas Bärgman, Daniel Nilsson, Bobbie Seppelt, Gustav Markkula, Giulio Bianchi Piccinini, and Trent Victor, *Great expectations: a predictive processing account of automobile driving*, Theoretical issues in ergonomics science **19** (2018), no. 2, 156–194.

[6] Harriet Feldman and Karl Friston, *Attention, uncertainty, and free-energy*, Frontiers in human neuroscience **4** (2010), 215.

[7] Karl Friston, Francesco Rigoli, Dimitri Ognibene, Christoph Mathys, Thomas Fitzgerald, and Giovanni Pezzulo, *Active inference and epistemic value*, Cognitive neuroscience **6** (2015), no. 4, 187–214.

[8] Karl J Friston, Jean Daunizeau, and Stefan J Kiebel, *Reinforcement learning or active inference?*, PloS one **4** (2009), no. 7, e6421.

[9] Ryota Kanai, Yutaka Komura, Stewart Shipp, and Karl Friston, *Cerebral hierarchies: predictive processing, precision and the pulvinar*, Philosophical Transactions of the Royal Society B: Biological Sciences **370** (2015), no. 1668, 20140169.

[10] Thomas Parr and Karl J Friston, *Generalised free energy and active inference*, Biological cybernetics **113** (2019), no. 5, 495–513.

[11] Steve M Potter, *What can ai get from neuroscience?*, 50 years of artificial intelligence, Springer, 2007, pp. 174–185.

[12] Madeleine Ransom and Sina Fazelpour, *The many faces of attention: Why precision optimization is not attention*, The Philosophy and Science of Predictive Processing (2020), 119.

[13] Madeleine Ransom, Sina Fazelpour, Jelena Markovic, James Kryklywy, Evan T Thompson, and Rebecca M Todd, *Affect-biased attention and predictive processing*, Cognition **203** (2020), 104370.

[14] Madeleine Ransom, Sina Fazelpour, and Christopher Mole, *Attention in the predictive mind*, Consciousness and cognition **47** (2017), 99–112.

[15] Oron Shagrir, *Why we view the brain as a computer*, Synthese **153** (2006), no. 3, 393–416.

[16] Rebecca M Todd, William A Cunningham, Adam K Anderson, and Evan Thompson, *Affect-biased attention as emotion regulation*, Trends in cognitive sciences **16** (2012), no. 7, 365–372.

Appendix A

Appendix

A.1 Code

All code used in the experiments can be found here:
`https://github.com/jortgutter/Thesis`

A.2 Plots

The following three pages have enlarged versions of the result plots.

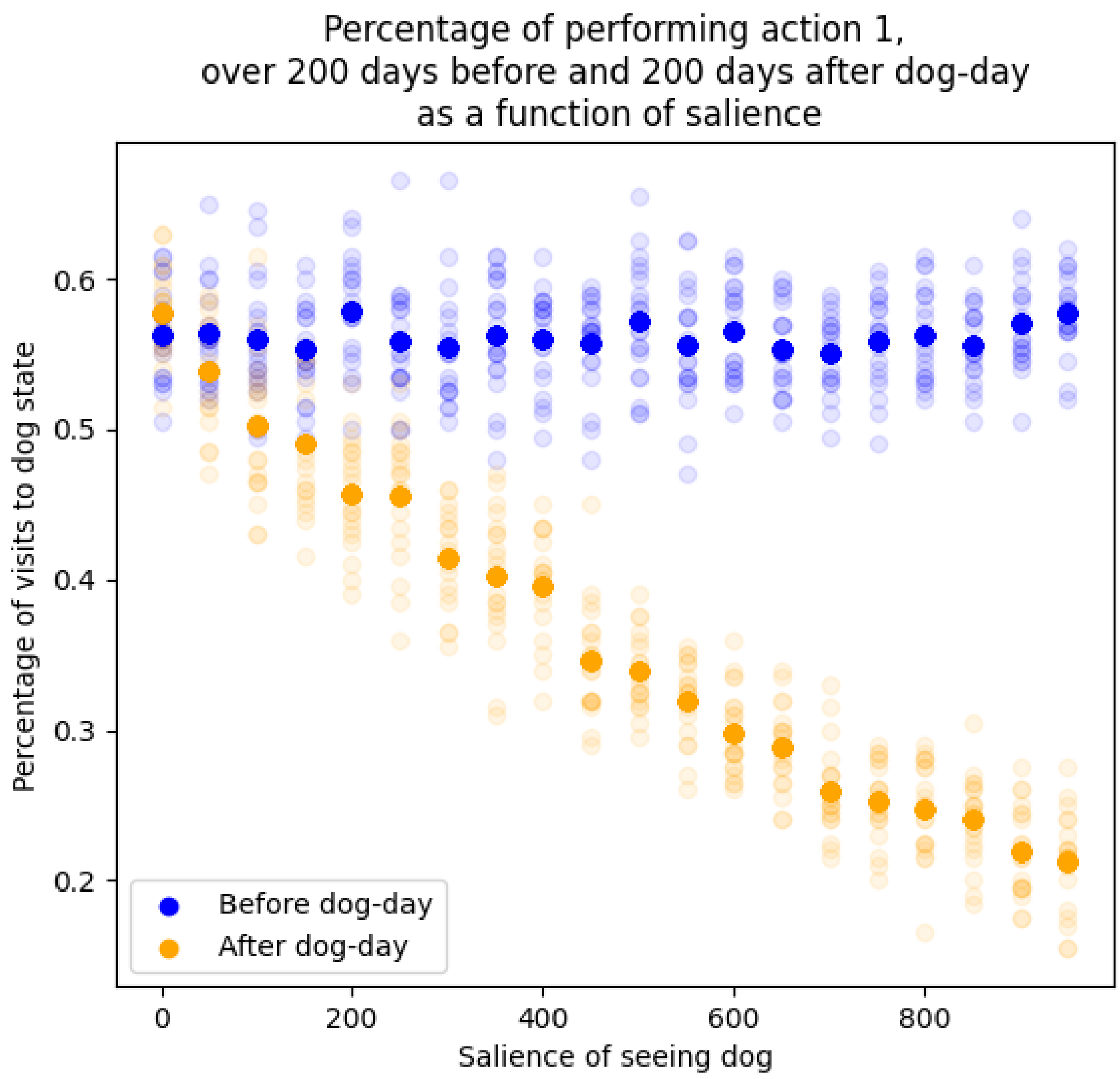

Figure A.1: Rotated version of the results of the varying salience experiment

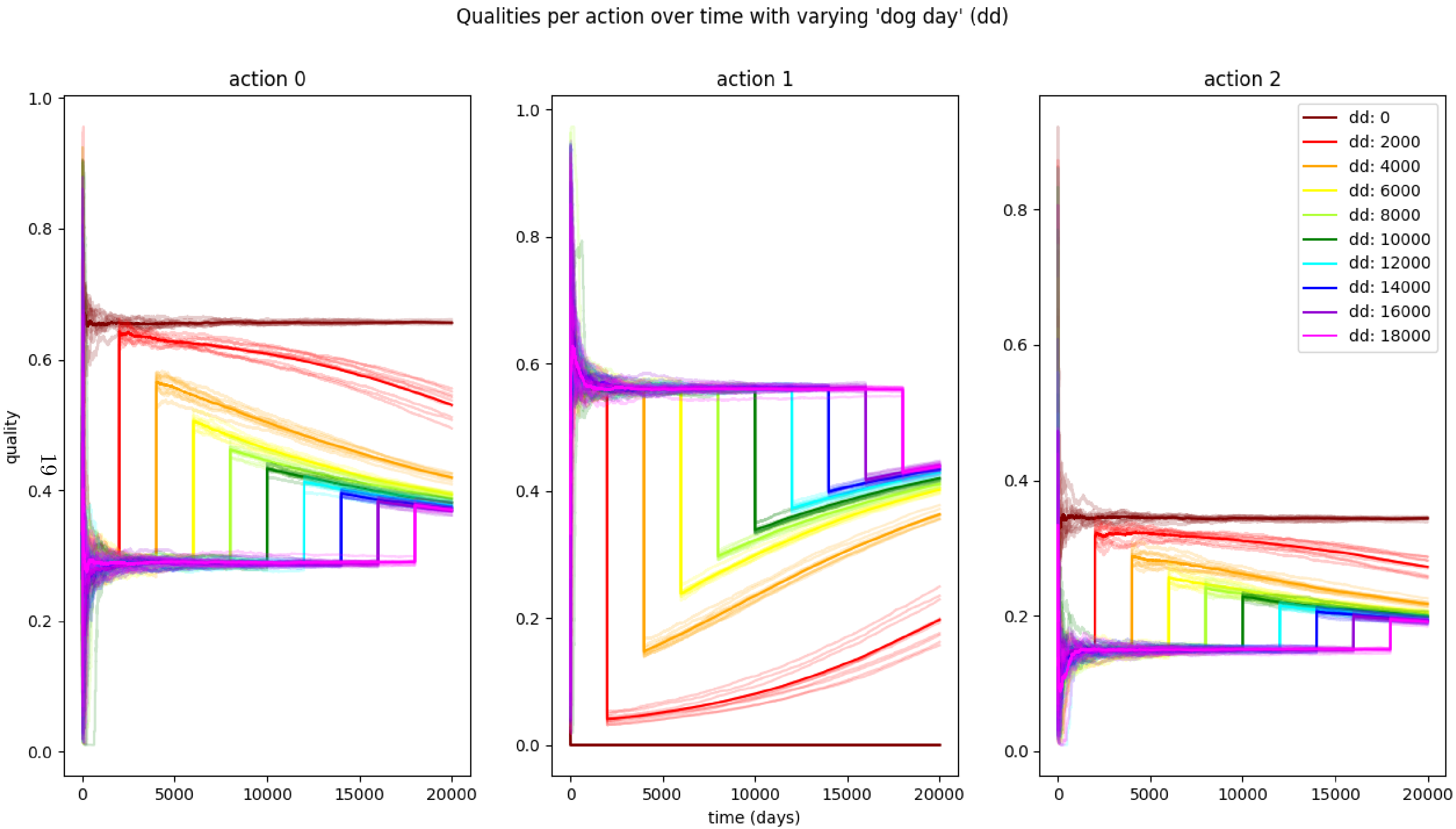

Figure A.2: Rotated version of the results of the varying dog day experiment

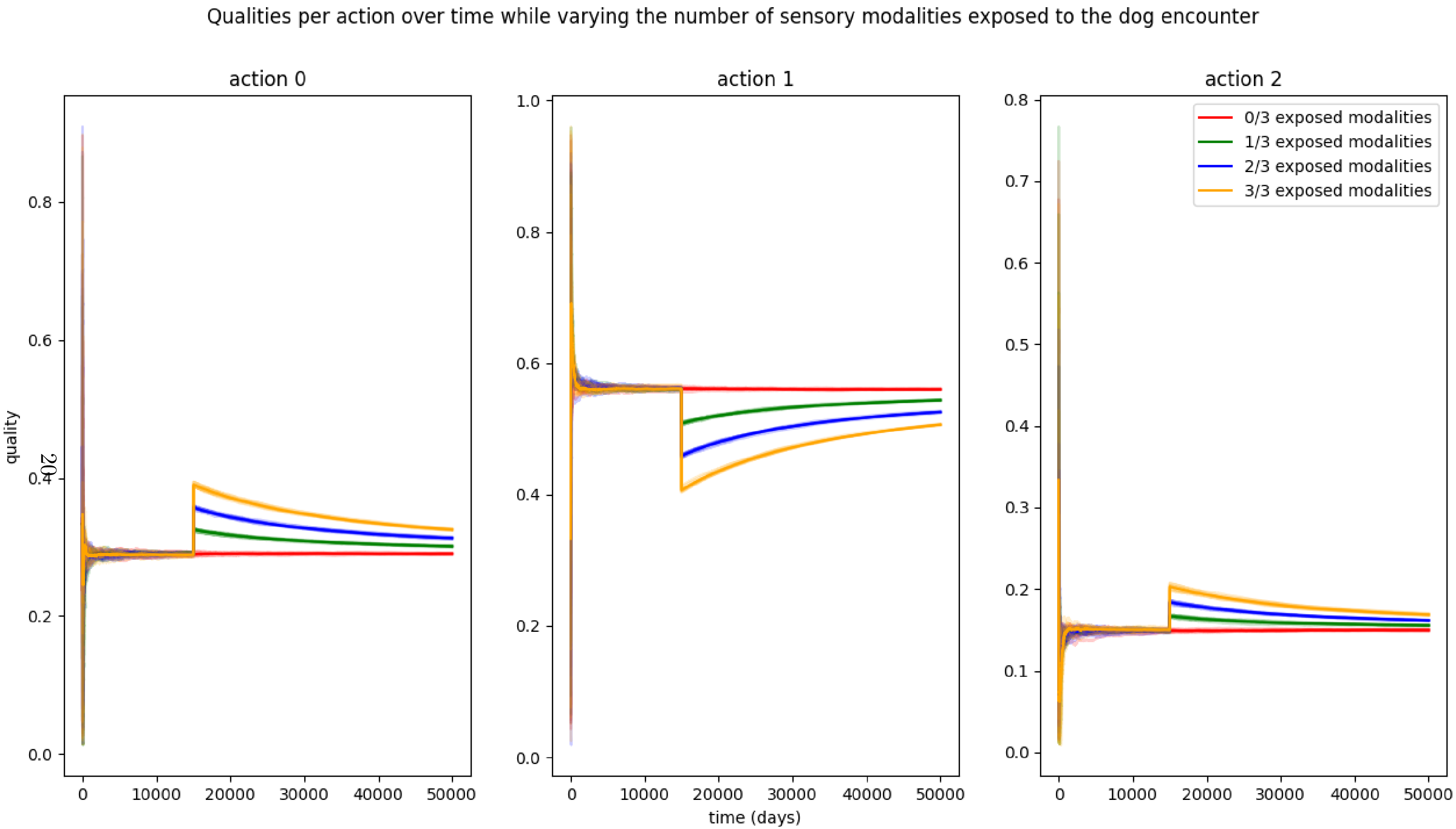

Figure A.3: Rotated version of the results of the varying dog modalities experiment