Adaptive Neurons Compute Confidence In A Decision Network

Luozheng Li
Peking University, Wangxuan Institute of Computer Technology
DaHui Wang (wangdh@bnu.edu.cn)
Beijing Normal University

Research Article

Keywords: substrate, neurons, artificial intelligence, adaptation

Posted Date: June 3rd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-558406/v1

Version of Record: A version of this preprint was published at Scientific Reports on November 12th, 2021. See the published version at https://doi.org/10.1038/s41598-021-01523-9.

Adaptive Neurons Compute Confidence in a Decision Network

Luozheng Li[1,3] **and DaHui Wang**[1,2,*]

[1]Beijing Normal University, School of Systems Science and State Key Lab of Cognitive Neuroscience and Learning, Beijing, 100875, China
[2]Beijing Normal University, Beijing Key Laboratory of Brain Imaging and Connectomics, Beijing, 100875, China
[3]Peking University, Wangxuan Institute of Computer Technology, Beijing, 100080, China
[*]wangdh@bnu.edu.cn

ABSTRACT

Human and many animals can asses the confidence of their decisions. However, little has been known about the underlying neural substrate and mechanism. Here we proposed a computational model consisting of one group of 'confidence neurons' with adaptation, which is able to asses the confidence of decision by detecting the slope of ramping activities of decision neurons. The simulated activities of 'confidence neurons' in our simple model capture the typical features of confidence in human and animals experiments. Our results indicate that confidence could be online formed along with the decision formation and the adaptation properties could be used to monitor the decision confidence in artificial intelligence.

Introduction

Confidence reflects subjective assessment of the choices[1,2]. In daily life, we frequently estimate confidence in our percepts and decisions. As a kind of metacognitive process, confidence implicates judgements about our own cognitive process[3–5]. However, the neural mechanism of confidence remains elusive.

A recent line of neurobiological studies identified the correlations between neural activities and decision confidence across species such as rats, monkeys and humans[5–11]. Experiments demonstrate that confidence may be encoded by some single neurons, according to a single source of uncertain evidence (one sensory modality or memory recalls). Furthermore, one recent research showed that single orbitofrontal cortex neurons in rats can encode abstract decision confidence irrespective of the sensory modality (olfactory or auditory). These studies provide solid biological basis for the neural representation of decision confidence. Nevertheless, it is still unclear how the confidence is computed during decision making in the neural circuits level.

Several theoretical models have tried to explore the confidence computations in the brain. Early researches modeled neural responses as probability distributions, where confidence is quantified by evaluating the posterior probability[12,13]. These models captured statistical characteristics of decision confidence but lack of neurobiological interpretability. Neural circuits model has been proposed to simulate the binary perceptual decision and the confidence was defined as the absolute difference between the firing rates of neuron population selective to the decision options[14]. This model successfully reproduced the observations in monkey experiment[15],but the neural circuit is difficult to calculate the absolute difference between firing rates of neuron pools. Some researchers assumed that decision was made by many loosely coupled modules each of which represent a stochastic sample of the sensory evidence integral and the confidence is encoded in the dispersion between modules[16]. However, the model did not explain how neural system read out the dispersion between modules.

In the present study, we constructed a neural circuit model to explain confidence formation during the decision process. The model consists of a classical decision module[17,18] and a confidence module. The confidence module receives the inputs from decision neurons. The activities of the confidence module can represent the abstract confidence observed in experiments. Different from previous theoretical studies, we introduce neural adaptation in the confidence module to achieve the computations biologically. Mathematically, adaptation enables the neurons to detect the slope of ramping activities of decision circuits. Thus, confidence computation and decision making can be implemented in one simple neural circuit.

Results

Model structure

The model consists of two parts: a classical decision circuit and a confidence module which includes recurrent connected neurons(as shown in Fig.1). The decision module has been well discussed in previous studies about two-alternative choice tasks[19–21]. It is composed of two groups of competing neurons (A and B), both groups receive feedforward inputs from

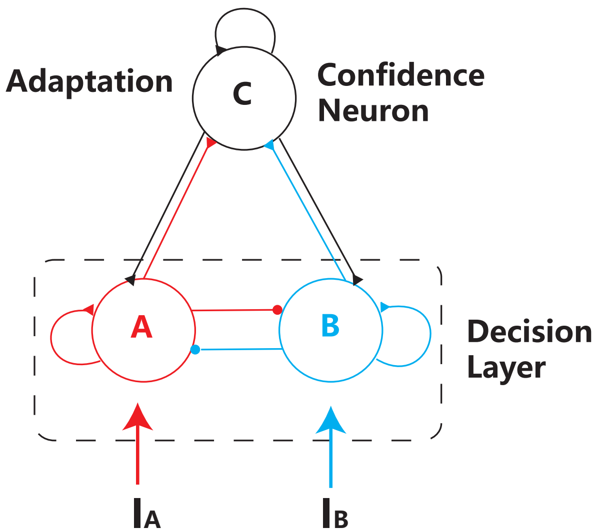


Figure 1. Model Structure. The model consists of decision layer and a confidence neuron pool. The decision layer follows classical decision circuit[17,18,21]. The adaptive confidence neurons receive the feedforward input from two competing neuron pools in decision layer and send the feedback projections to the decision layer.

upstream neurons and feedback currents from the confidence neurons. The confidence module (C) consists of one group of neurons whose activities reflect the confidence of decisions. Neurons in the confidence module are innervated by both neural groups in the decision circuit(A and B) and send feedba als to them. Since that confidence is the evaluation of the decisions, the confidence neurons equally receive the information from any decision neurons.

Ramping activities in the decision module

In the decision module, neuron firing rates displayed ramping activity during the stimulus presentation($0-1000ms$ in Fig2), which is consistent with previous electrophysiological[22] and theoretical studies[17,18]. Larger value of c leads to steeper ramping activity (as shown in Fig2) and shorter decision time. At the same time, the larger c and shorter decision time implies an easier task and subject should have higher confidence in experiment. Thus, the slope of ramping activity can be thought as a signal of confidence of the decision. If the downstream of neural circuit can detect the slope of ramping activity, the confidence signal can be read out.

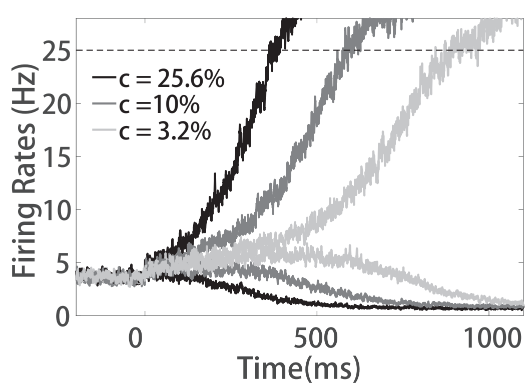


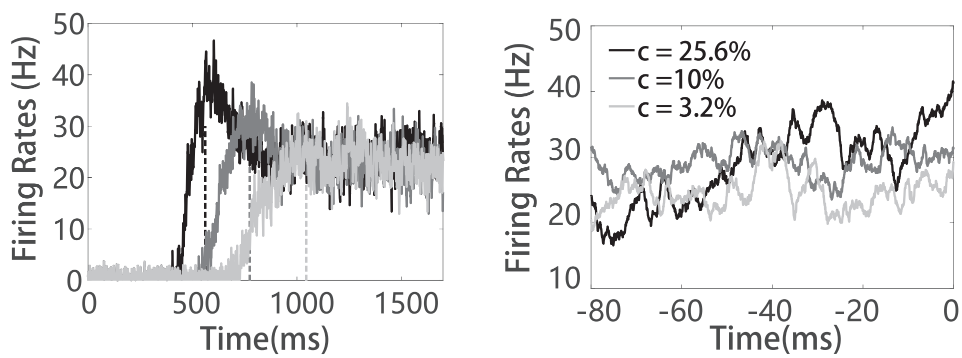
Figure 2. Ramping activities of the neurons in decision layer during Decision making. The dash line indicates the decision threshold.

Activities of confidence neurons

In our model, confidence neurons are designed to detect the slope of ramping activity of decision neurons with adaptation mechanism. The neurons receive excitatory feedforward inputs from decision layer, as well as inhibitory currents caused by

adaptation. Based on biological evidence, adaptive currents will increase with the firing rates of confidence neuron with some time delay (see Eq.9). Thus, the activities of confidence neurons first increase along with the ramping activities of decision layer and then elicit inhibitions caused by adaptive currents. When the stimulus is easy to be discriminated, ramping activities in decision layer have a large slope (see Fig2). Because of the large time constant of adaptive currents(τ_a in Eq.9), the inhibitory adaptive currents($a(t)$) cannot keep pace with the increasing inputs caused by rapid ramping activities($r_{in}(t)$). As the result, with the integration of time(see Eq.7), the confidence neurons receive weak inhibitions caused by adaptation and reach large firing rates at the decision moment. On the contrary, when a difficult task is given, the inhibitory adaptive currents ($a(t)$) can catch up with the inputs from ramping up activities ($r_{in}(t)$), which leads to larger inhibitory adaptive currents and lower firing rates at the decision moment.

In the simulations, we define the decision moment following the convention of the previous study[18]: once the ramping activity of decision circuit exceeds the decision threshold (25Hz), the network makes a choice. Fig3a shows a sample trial of confidence neurons' activities. Firing rates of confidence neurons ramp up to different levels as the difficulty of the task (c). Steeper ramping activities of decision neuron (larger c) corresponds to the higher firing rates of confidence neuron at the decision moment. For clarity, we also align the time of confidence neurons' firing rates to the decision moment (Fig3b). With more precise time scales, we can clearly see that the activities of confidence neuron are negatively correlated with the task difficulty at the decision time (Fig3b).



(a) Neural activities aligned to the stimulus onset. **(b)** Neural activities aligned to the decision moment.

Figure 3. Activities of confidence neurons during decision making. The vertical dash line in (a) indicate the decision moment. The horizontal axis in (b) shows the time to the decision moment.

Typical features of reports by confidence neurons

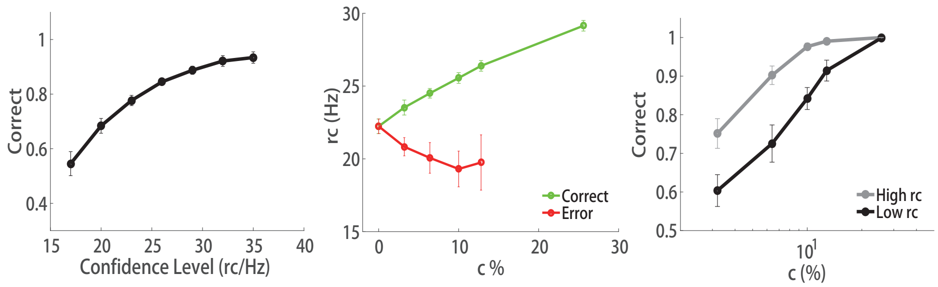
Activities of confidence neurons may be affected by noise in a single trial, so it is necessary to analysis its statistical behaviors. In the simulation, the confidence report or neural representation of confidence, rc, is depicted by the mean firing rates of confidence neurons in the interval of $10ms$ just before the decision moment. Simulation results reveal that the statistical behaviors of rc (Fig. 4a) are consistent with the typical features of the general confidence reported in human and animal experiments[1,5,23–25] (Fig4b):

- The decision accuracy is positively correlated with the confidence level (indirect measurement or direct report in experiment) (Left panel in Fig4b).
- By splitting the trials into correct and error trials, it can be found that the confidence level of trials with correct decisions will increase as the task difficulty decreases, and the opposite results is obtained on the error trials (Middle panel in Fig4b).
- The psychometric curve of trials that report high confidence will shift upward (Right panel in Fig4b).

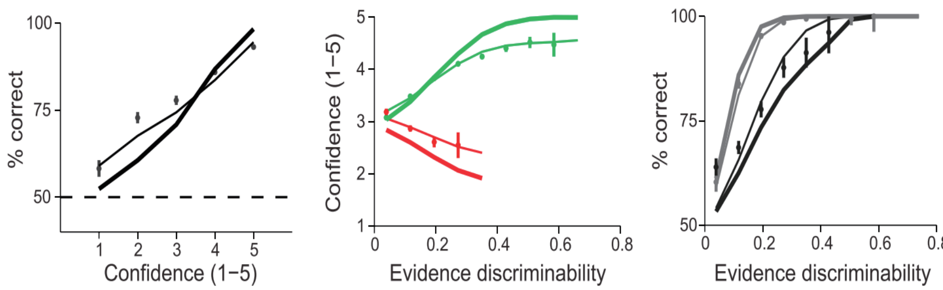
In brief, activities of the confidence neurons in our circuit model behave as the general confidence observed in experiments. These results suggest that confidence could be formed along with decision and represented by a separate group of neurons.

Sensitivity of adaptation on confidence coding

Adaptation of confidence neurons plays key roles in the confidence formation during decision process. In our model, neural adaptation is described by two parameters (in Eq9), A_1, reflecting the strength of the adaptation caused by spikes, and A_0 indicating the baseline level of adaptation in the resting state. Since different adaptation strength causes different firing rates



(a) r_c as the confidence report in the simulations



(b) Confidence measured in the human behavioral experiment

Figure 4. Simulated confidence. (a) The reported confidence (rc) by confidence neuron are consistent with the reported confidence in human experiments. Error bars show the standard error of 10 sessions. (b) The reported confidence in human experiments, adapted from experimental study[1].

of the confidence neurons at random level($\%c = 0$), we use 'relative rc'($rc - rc(\%c = 0)$) instead of rc to show the changes of slopes. To investigate the influence of adaptation parameters on the confidence coding, we calculate 'relative rc' over the different coherence levels given varied A_1 and A_0. In Fig5a, we plot the curves with different A_1 and fixed A_0, and similar curves are shown in in Fig5b given different A_0 and fixed A_1.

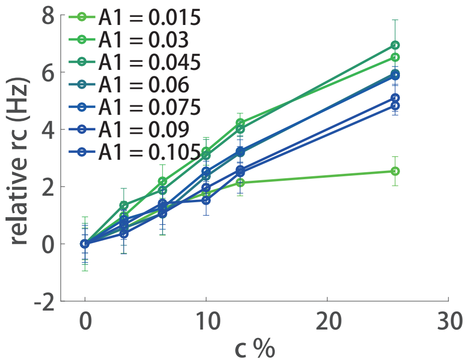
The slopes of the curves in Fig5a and Fig5b reflect the coding ability of confidence neurons. The slope equals to zero means the confidence neurons have same firing rates given stimuli with different coherence levels, which implies confidence neurons can not code the decision confidence. Larger slope indicates larger difference of the firing rates between stimuli with different coherence level, which means confidence neurons are more sensitive to the change of confidence. Fig5c shows the dependence of the slope on parameter A_1 and A_0, where horizontal dash line is shown in Fig5a and the vertical line is shown in Fig5b. These results indicate that adaptation modulation is statistically robust, because parameters in a large range (yellow areas in Fig5c) support the confidence coding.

Discussion

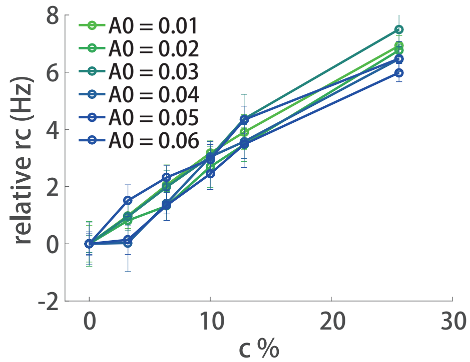
In the present study, we proposed a computational model in which the decision confidence can be computed and represented in a simple neural circuit. We suggested that the confidence representation can be achieved by neural adaptation, which provides common negative feedbacks in neural system. Based on the previous observations in experiments[2,5] and theoretical models of decision making[17,18,21], we designed the confidence neurons as a neural group whose activaty reflects the confidence level of the decisions. Simulation results confirmed that the activities of confidence neurons successfully capture the general features of confidence consistently revealed in animals and human behavioral experiments[1,5,23,24]. At last, we explored the influence of adaptation parameters on the confidence coding, and demonstrated the adaptation modulation is statistically robust.

In the study, we proposed one group of confidence neurons. Number of evidences support the exist of confidence neurons in brain. First, a recent experimental identified that single orbitofrontal cortex neuron in rats can encode general decision confidence[25]. Second, some neurons in the rats orbitofrontal cortex area reflect uncertainty during decision making[2], which has been modeled by the reciprocal connected circuits[16]. Third, single-neuron in human medial temporal lobe has been found signaling the confidence during decision making[11] At last, based on our model, confidence neuron should have neural ramping activities without preferred directions and this kind of activity was observed in monkeys' LIP.

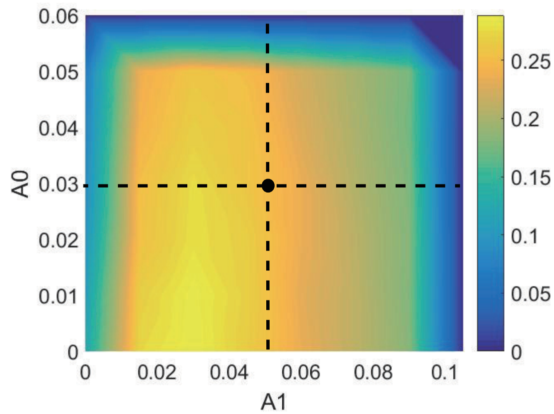
A debated question is whether the neural system need to represent confidence in circuits, since the system can utilize



(a) Relative r_C with varied A1



(b) Relative r_C with varied A0



(c) Coding capacity over adaptation parameters

Figure 5. The influence of adaptation parameters on confidence representation. (a) The effects of $A1$; (b) The effects of $A0$; (c) The dependence of coding capacity on adaptation parameters.The error bars in (a) and (b) show the standard error of 10 sessions. Different colors in (c) code the average slope of rc.

decoding algorithms to read out confidence from decision layer directly. For instance, researchers believed that neural system may encode the confidence in form of reaction time[5]. However, experiments showed the reaction time heuristic cannot fully account for confidence reports[1]. And behavioral experiments show that choice accuracy and confidence reporting are separated processes[5,15,26], suggesting that confidence computation may not be accomplished in decision layer. As our model proposed, decision confidence is computed in neural circuits without extra decoding strategy, which is more simplicity in theoretic and reasonable in biology.

Methods

Dynamics of the decision circuit

Spiking neuron models[17]and reduced mean-field models[18] have been proposed in the previous theoretical studies to explain the mechanisms underlying binary decisions. The spiking model is more biological while the mean-filed model is concise and convenient for theoretical analysis. Both kinds of model successfully replicated the most of the psychophysical and physiological results in the monkey experiments[18,21]. Here we adopt the mean-field model to describe the neural dynamics in decision circuit. As described in previous work[17,18,21], dynamics of the neurons in decision module can be described by the slow dynamics of N-methyl-D-aspartic acid(NMDA) receptors:

$$\frac{dS_i}{dt} = -\frac{S_i}{\tau_{NMDA}} + (1 - S_i)\gamma r_i, \tag{1}$$

where S_i is the gating variable of NMDA, i is A or B, standing for the group label. τ_{NMDA} is the decay time constant of NMDA. γ is a constant that controls the strength of the gain of S_i caused by firing rates. r_i represents the firing rates of the two neural population. The dynamics of r_i is given by:

$$r_i = \phi(I_{syn,i}), \tag{2}$$

$$I_{syn,i} = J_{ii}S_i - J_{ij}S_j + I_0 + I_{ext,i} + I_{noise,i}, \tag{3}$$

where $\phi(x)$ is the input-output function of single neuron, describing the relation between synaptic input currents and neural firing rates. $I_{syn,i}$ represents the synaptic currents of neural group i(A or B). J_{ii} and J_{ij} are the strength of recurrent connections and cross inhibition, respectively. I_0 is the background input without bias, while I_i is the stimulus to the population i with varied strength. $I_{noise,i}$ is a noise term.

As in the previous work[18], function $\phi(x)$ is chosen as:

$$\phi(I) = \frac{c_E I - I_{th}}{1 - \exp[-g_E(c_E I - I_{th})]}, \tag{4}$$

where c_E is the gain factor, I_{th} is the threshold current, g_E is a noise factor determining the nonlinearity of the function.

Dynamics of confidence neurons

We consider a group of neurons that receive inputs from the decision circuit. It is named 'confidence neurons' since we can read out the confidence of decision according to its activities. The dynamics of confidence neuron is similar to the neurons in decision module, except for the adaptation currents,

$$\frac{dr_C}{dt} = -\frac{r_C}{\tau_r} + \phi_C(I_{syn,C}), \tag{5}$$

$$\frac{dS_C}{dt} = \frac{-S_C}{\tau_{NMDA}} + \gamma(1 - S_C)r_C, \tag{6}$$

$$I_{syn,C} = J_C S_C + J_{dc} r_{in} + I_{0c} - J_a a + I_{noise,C}, \tag{7}$$

where r_C is the firing rate of the confidence neurons, τ_r the time constant of firing rates, usually $2-5ms$. ϕ_C descried the Input-Output function of the confidence neuron, which is simplified as:

$$\phi_C(I) = \max(c_E I - I_{th,C}, 0), \tag{8}$$

$I_{syn,C}$ is the synaptic currents, and S_C is the gating variable of NMDA. J_C denotes the strength of the recurrent connection between confidence neurons. $r_{in} = r_A + r_B$, indicates the inputs from decision layer. J_{dc} is the connection strength from decision layer to confidence neurons. Adaptation currents are denoted by a and controlled by constant J_a.

Adaptation is very common in the nervous system. Previous studies revealed that many cellular mechanisms could contribute to the neural adaptation. These mechanisms can be divided into two classes[27,28]: the spike-triggered mechanism, e.g., the calcium-activated potassium current, and the subthreshold voltage-dependent mechanism, e.g., the voltage-gated potassium current. Here we model adaptation currents based on these two general mechanisms and the adaptation current of the confidence neuron is given by:

$$\frac{da}{dt} = -\frac{a}{\tau_a} + A_1 r_C + A_0, \tag{9}$$

where τ_a is the time constant of adaptation and reflects the slow dynamics of calcium currents. Parameter A_1 denotes the strength of adaptation caused by spikes, while A_0 is the strength of subthreshold adaptation.

Simulation protocol

To explain the confidence computations, we simulate the general two-alternative decision-making task with reaction-time style. Many similar experiments have on performed on monkeys[22,29] and humans[1].

Second order Runge-Kutta method with an time step of 0.05 ms is applied for numerical simulations. Parameters in the simulations are chosen as in Table1 without specification.

In a single trial, we simulate the model for a fixed time period $T = 1,500$ ms. The network receives only unbiased background inputs from $t = -200$ ms to 0 ms. Biased stimulus is onset at $t = 0$ ms, and decision circuit receives external inputs from $t = 0$ ms to $t = 1,000$ ms, The stimulus is set as biased inputs as in[18]:

$$I_{ext} = J_{ext}\mu_0(1 \pm \frac{c}{100\%}), \tag{10}$$

where c stands for the task difficulty, which is the coherence level in a dot-motion task[22], larger c corresponds to an easier trials. J_{ext} is the average synaptic coupling with AMPAR receptors, μ_0 stands for the absolute stimulus strength. Decision is made when the firing rates of the two competing neural groups reaches a threshold ($\theta = 25Hz$).

To compare with the experimental results, we calculated the average value of r_C across a time of $10ms$ before the decision time as the indicator of confidence. To investigate the statistic features of activities of confidence neurons, we employ 10 sessions, with 500 trials in each session. For the simulations of each value of adaptation parameters, we also employ 10 sessions, with 500 trials in each session.

Parameter	value
τ_{NMDA}, time constant of NMDA receptors	0.1s
τ_a, time constant of adaptation	0.25 s
τ_r, time constant of firing rate	0.002 s
θ, decision threshold	25 Hz
γ, NMDA gain factor per spike	0.641
J_{ii}, synaptic strength within neural groups	0.2609 nA
J_{ij}, synaptic strength between neural groups	0.0497 nA
J_C, synaptic strength between confidence neurons	0.15 nA
J_{ext}, external input synaptic strength to decision layer	0.15 nA
J_{dc}, feedforward synaptic strength	0.015 nA/Hz
J_a, gain of adaptive currents to confidence neuron	0.001
c_E, slope of the F-I function of decision neurons	270/(VnC)
I_{th}, firing threshold of decision neurons	108 Hz
g_E, noise factor of decision neurons	0.154 s
$I_{th,C}$, threshold of confidence neurons	108 Hz
A_1, strength of adaptation caused by spikes,	0.05 nA
A_0, strength of subthreshold adaptation	0.03 nA·Hz
μ_0, average external inputs	30 Hz
I_0, background inputs in decision layer	0.3255 nA
I_{0c}, background inputs in confidence neurons	0.2 nA

Table 1. Parameters used in the model

References

1. Sanders, J. I., Hangya, B. & Kepecs, A. Signatures of a statistical computation in the human sense of confidence. *Neuron* **90**, 499–506, DOI: https://doi.org/10.1016/j.neuron.2016.03.025 (2016).
2. Kepecs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231, DOI: https://doi.org/10.1038/nature07200 (2008).
3. Charles, L., Van Opstal, F., Marti, S. & Dehaene, S. Distinct brain mechanisms for conscious versus subliminal error detection. *Neuroimage* **73**, 80–94, DOI: https://doi.org/10.1016/j.neuroimage.2013.01.054 (2013).
4. Flavell, J. H. Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *Am. psychologist* **34**, 906, DOI: https://doi.org/10.1037/0003-066X.34.10.906 (1979).
5. Lak, A. *et al.* Orbitofrontal cortex is required for optimal waiting based on decision confidence. *Neuron* **84**, 190–201, DOI: https://doi.org/10.1016/j.neuron.2014.08.039 (2014).
6. Lau, H. C. & Passingham, R. E. Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proc. Natl. Acad. Sci.* **103**, 18763–18768, DOI: https://doi.org/10.1073/pnas.0607716103 (2006).
7. Fleming, S. M. & Dolan, R. J. Effects of loss aversion on post-decision wagering: Implications for measures of awareness. *Conscious. cognition* **19**, 352–363, DOI: https://doi.org/10.1016/j.concog.2009.11.002 (2010).
8. De Martino, B., Fleming, S. M., Garrett, N. & Dolan, R. J. Confidence in value-based choice. *Nat. neuroscience* **16**, 105, DOI: https://doi.org/10.1038/nn.3279 (2013).
9. Fleck, M. S., Daselaar, S. M., Dobbins, I. G. & Cabeza, R. Role of prefrontal and anterior cingulate regions in decision-making processes shared by memory and nonmemory tasks. *Cereb. Cortex* **16**, 1623–1630, DOI: https://doi.org/10.1093/cercor/bhj097 (2006).

10. Kim, H. & Cabeza, R. Trusting our memories: dissociating the neural correlates of confidence in veridical versus illusory memories. *J. Neurosci.* **27**, 12190–12197, DOI: https://doi.org/10.1523/JNEUROSCI.3408-07.2007 (2007).

11. Unruh-Pinheiro, A. *et al.* Single-neuron correlates of decision confidence in the human medial temporal lobe. *Curr. Biol.* **30**, 4722–4732, DOI: https://doi.org/10.1016/j.cub.2020.09.021 (2020).

12. Beck, J. M. *et al.* Probabilistic population codes for bayesian decision making. *Neuron* **60**, 1142–1152, DOI: https://doi.org/10.1016/j.neuron.2008.09.021 (2008).

13. Fiser, J., Berkes, P., Orbán, G. & Lengyel, M. Statistically optimal perception and learning: from behavior to neural representations. *Trends cognitive sciences* **14**, 119–130, DOI: https://doi.org/10.1016/j.tics.2010.01.003 (2010).

14. Wei, Z. & Wang, X.-J. Confidence estimation as a stochastic process in a neurodynamical system of decision making. *J. neurophysiology* **114**, 99–113, DOI: https://doi.org/10.1152/jn.00793.2014 (2015).

15. Kiani, R. & Shadlen, M. N. Representation of confidence associated with a decision by neurons in the parietal cortex. *science* **324**, 759–764, DOI: DOI:10.1126/science.1169405 (2009).

16. Paz, L., Insabato, A., Zylberberg, A., Deco, G. & Sigman, M. Confidence through consensus: a neural mechanism for uncertainty monitoring. *Sci. reports* **6**, 21830, DOI: https://doi.org/10.1038/srep21830 (2016).

17. Wang, X.-J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968, DOI: https://doi.org/10.1016/S0896-6273(02)01092-9 (2002).

18. Wong, K.-F. & Wang, X.-J. A recurrent network mechanism of time integration in perceptual decisions. *J. Neurosci.* **26**, 1314–1328, DOI: https://doi.org/10.1523/JNEUROSCI.3733-05.2006 (2006).

19. Cutsuridis, V., Kahramanoglou, I., Smyrnis, N., Evdokimidis, I. & Perantonis, S. A biophysical neural accumulator model of decision making in an antisaccade task. *Neurocomputing* **70**, 1390–1402, DOI: https://doi.org/10.1016/j.neucom.2006.06.002 (2007). Advances in Computational Intelligence and Learning.

20. Martí, D., Deco, G., Giudice, P. D. & Mattia, M. Reward-biased probabilistic decision-making: Mean-field predictions and spiking simulations. *Neurocomputing* **69**, 1175–1178, DOI: https://doi.org/10.1016/j.neucom.2005.12.069 (2006). Computational Neuroscience: Trends in Research 2006.

21. Wang, X.-J. Decision making in recurrent neuronal circuits. *Neuron* **60**, 215–234, DOI: https://doi.org/10.1016/j.neuron.2008.09.034 (2008).

22. Roitman, J. D. & Shadlen, M. N. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. neuroscience* **22**, 9475–9489, DOI: https://doi.org/10.1523/JNEUROSCI.22-21-09475.2002 (2002).

23. Lak, A., Nomoto, K., Keramati, M., Sakagami, M. & Kepecs, A. Midbrain dopamine neurons signal belief in choice accuracy during a perceptual decision. *Curr. Biol.* **27**, 821–832, DOI: https://doi.org/10.1016/j.cub.2017.02.026 (2017).

24. Navajas, J. *et al.* The idiosyncratic nature of confidence. *Nat. human behaviour* **1**, 810–818, DOI: https://doi.org/10.1038/s41562-017-0215-1 (2017).

25. Masset, P., Ott, T., Lak, A., Hirokawa, J. & Kepecs, A. Behavior-and modality-general representation of confidence in orbitofrontal cortex. *Cell* DOI: https://doi.org/10.1016/j.cell.2020.05.022 (2020).

26. Higham, P. A. No special k! a signal detection framework for the strategic regulation of memory accuracy. *J. Exp. Psychol. Gen.* **136**, 1, DOI: https://doi.org/10.1037/0096-3445.136.1.1 (2007).

27. Benda, J. & Herz, A. V. A universal model for spike-frequency adaptation. *Neural computation* **15**, 2523–2564, DOI: https://doi.org/10.1162/089976603322385063 (2003).

28. Brette, R. & Gerstner, W. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J. neurophysiology* **94**, 3637–3642, DOI: https://doi.org/10.1152/jn.00686.2005 (2005).

29. Furman, M. & Wang, X.-J. Similarity effect and optimal control of multiple-choice decision making. *Neuron* **60**, 1153–1168, DOI: https://doi.org/10.1016/j.neuron.2008.12.003 (2008).

Acknowledgements

This work was supported by National Key R&D Program of China [grant numbers: 2019YFA0709503]

Author contributions statement

D.W. and L.L contributed to the model design. L.L performed the simulations and data fitting. D.W. and L.L wrote and reviewed the manuscript.

Additional information

Competing interests: The authors declare no competing financial interests.

Figures

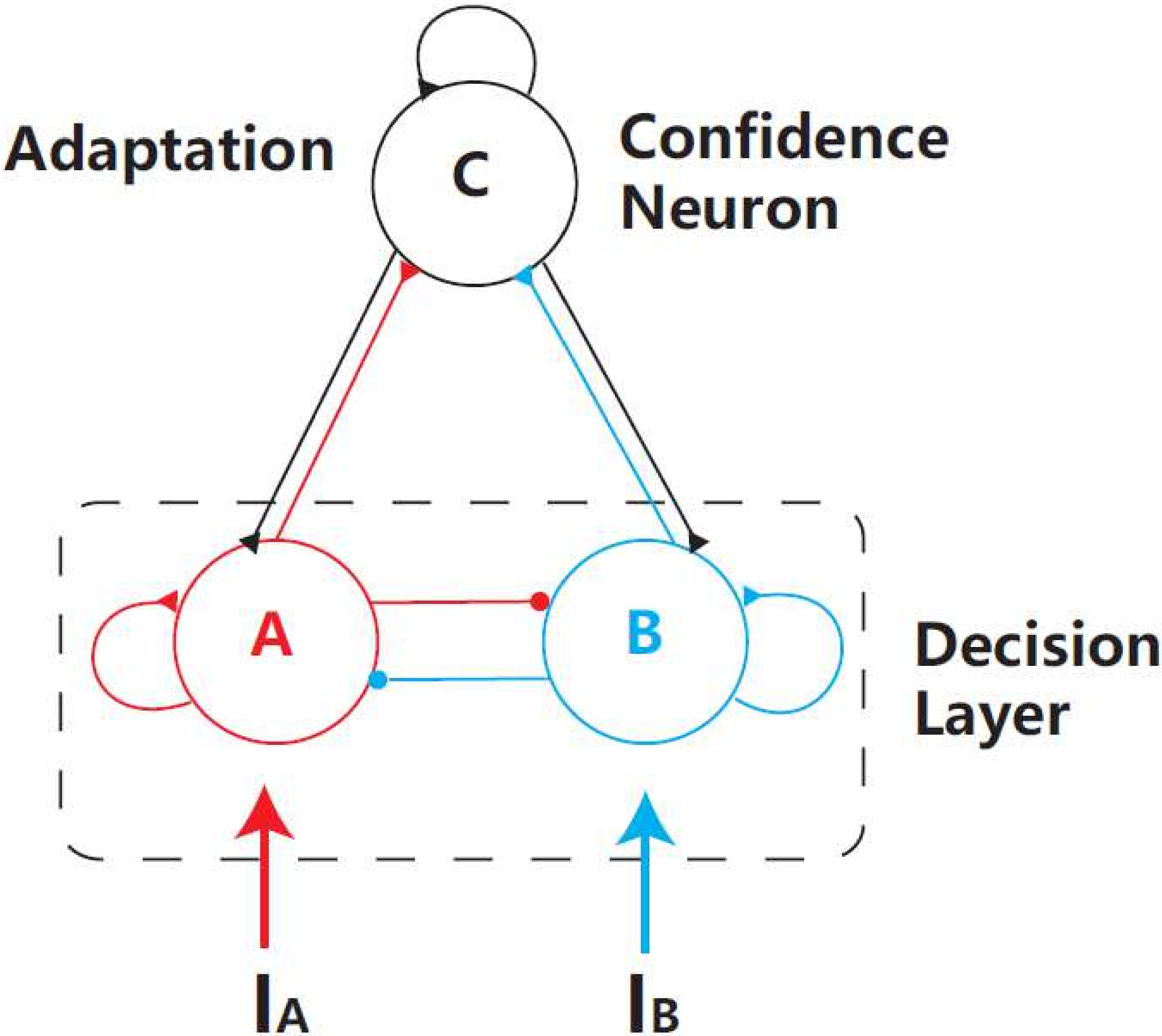


Figure 1

Model Structure. The model consists of decision layer and a confidence neuron pool. The decision layer follows classical decision circuit17, 18, 21. The adaptive confidence neurons receive the feedforward input from two competing neuron pools in decision layer and send the feedback projections to the decision layer.

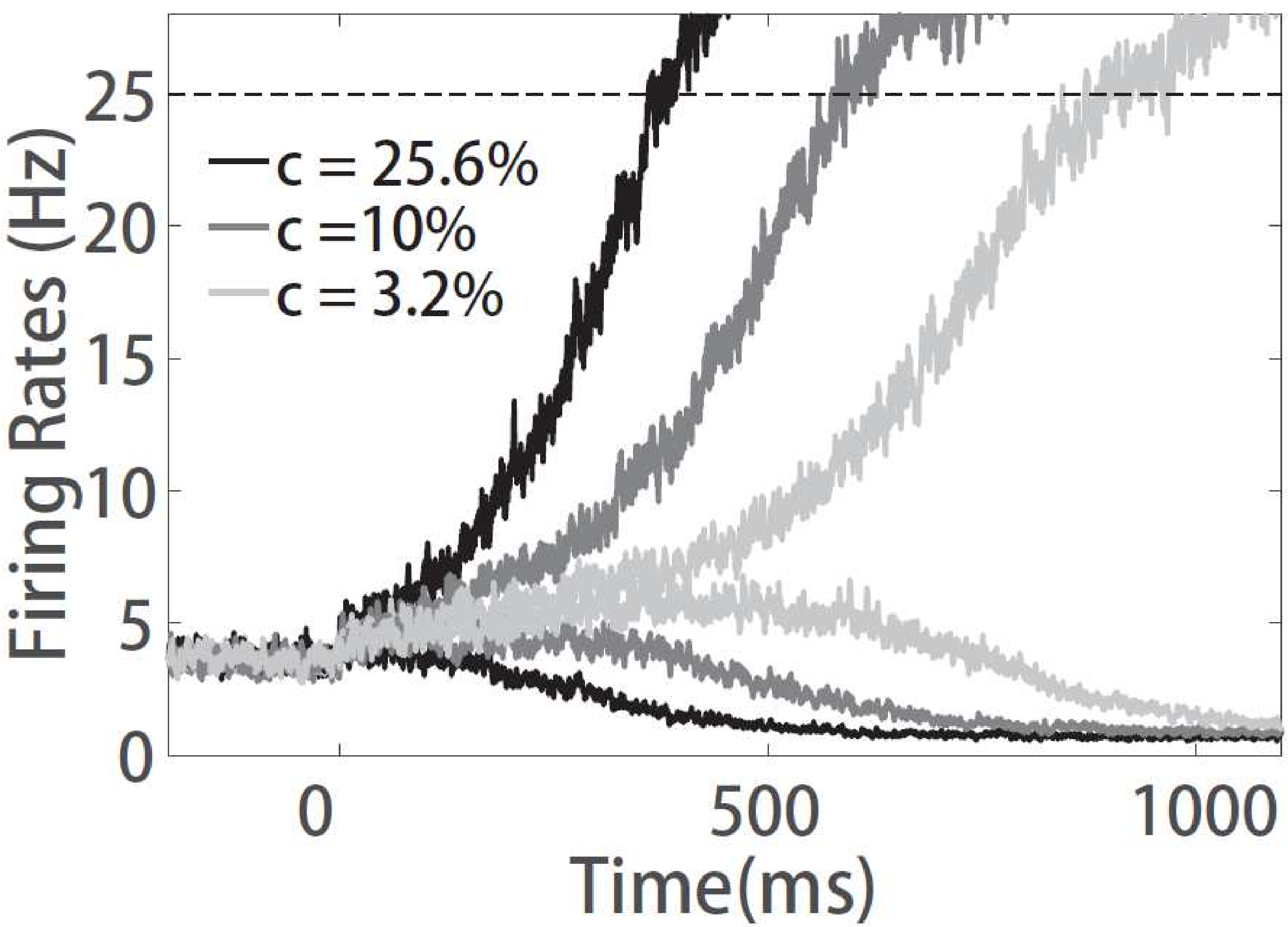
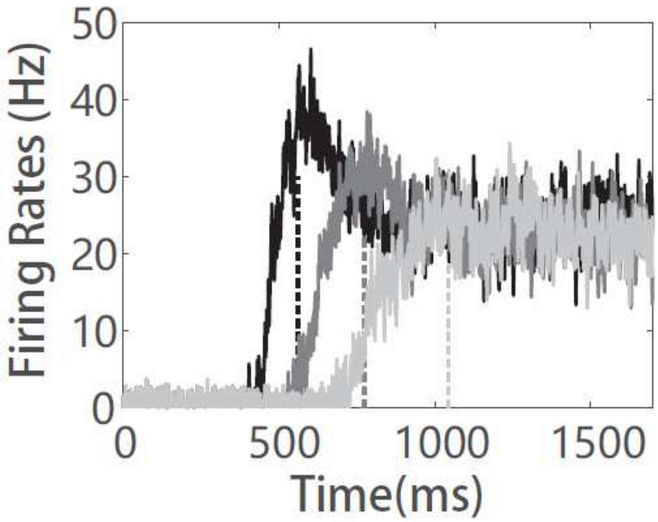
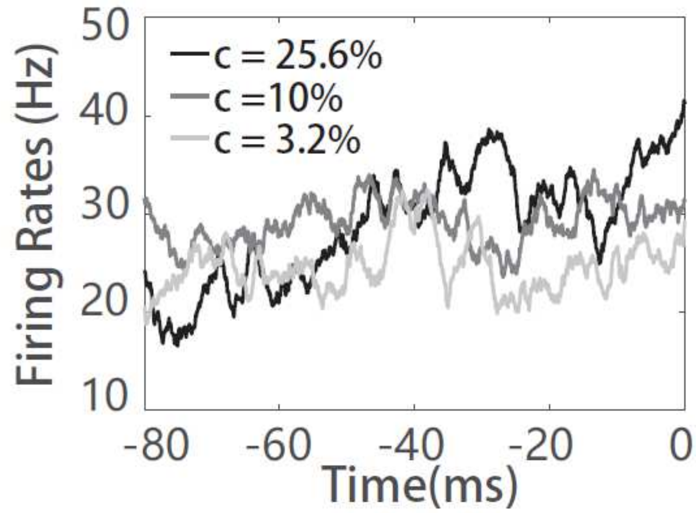


Figure 2

Ramping activities of the neurons in decision layer during Decision making. The dash line indicates the decision threshold.



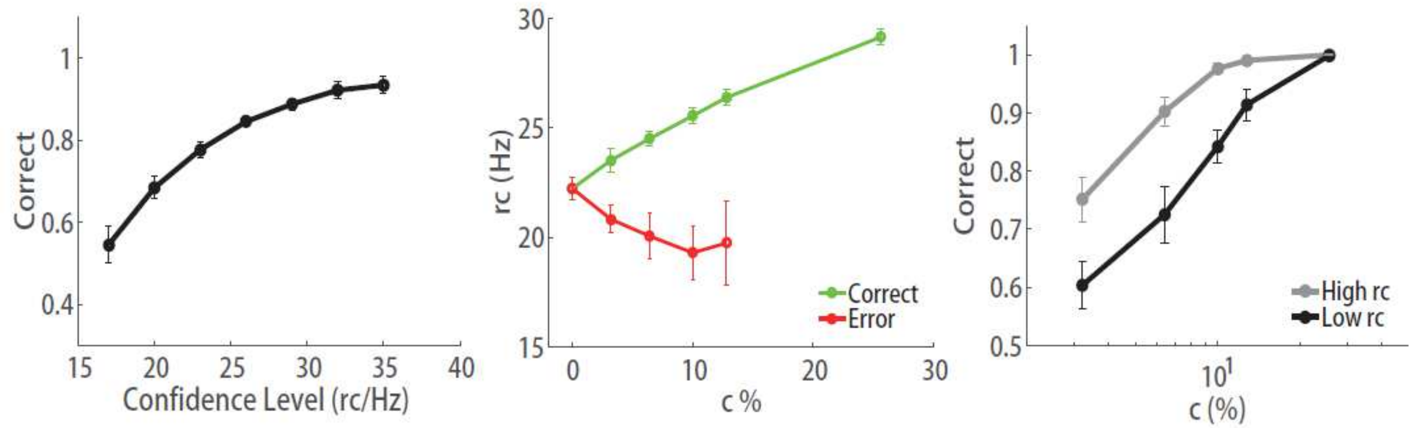
(a) Neural activities aligned to the stimulus onset.



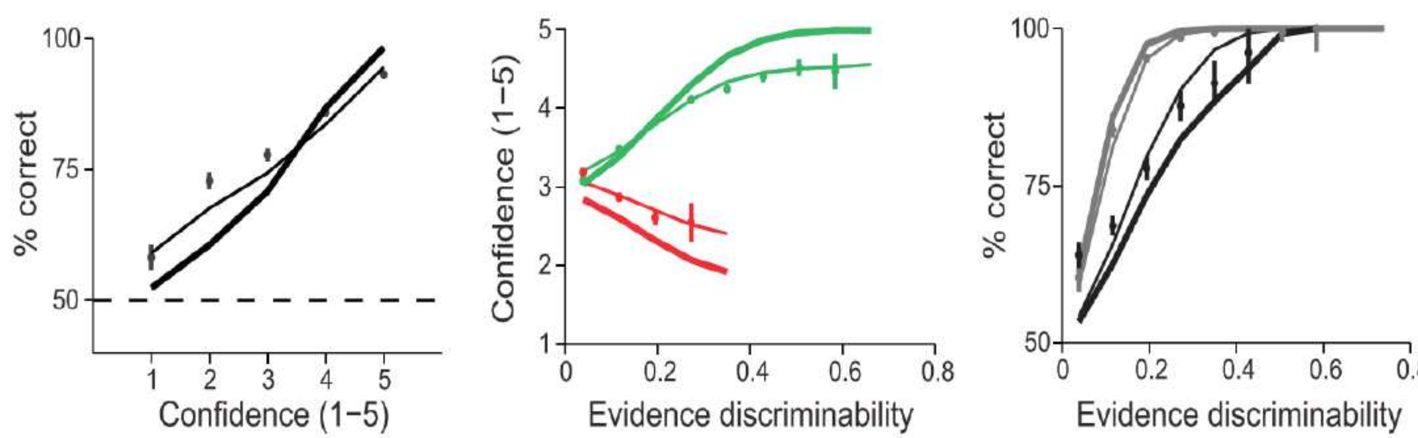
(b) Neural activities aligned to the decision moment.

Figure 3

Activities of confidence neurons during decision making. The vertical dash line in (a) indicate the decision moment. The horizontal axis in (b) shows the time to the decision moment.



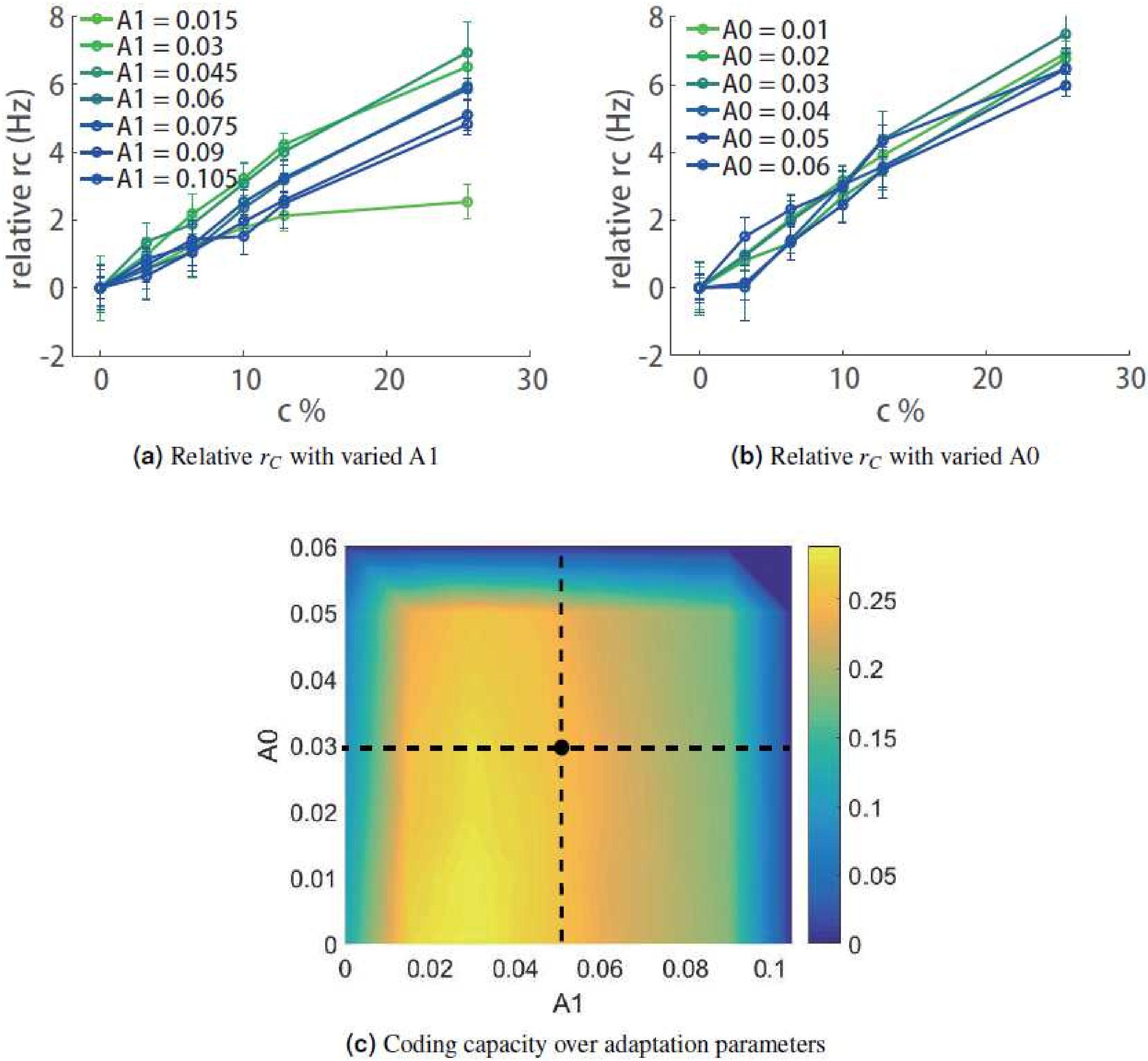
(a) r_c as the confidence report in the simulations



(b) Confidence measured in the human behavioral experiment

Figure 4

Simulated confidence. (a) The reported confidence (rc) by confidence neuron are consistent with the reported confidence in human experiments. Error bars show the standard error of 10 sessions. (b) The reported confidence in human experiments, adapted from experimental study1.



(a) Relative r_C with varied A1

(b) Relative r_C with varied A0

(c) Coding capacity over adaptation parameters

Figure 5

The influence of adaptation parameters on confidence representation. (a) The effects of A1; (b) The effects of A0; (c) The dependence of coding capacity on adaptation parameters.The error bars in (a) and (b) show the standard error of 10 sessions. Different colors in (c) code the average slope of rc.