Behavioral and pharmacological evidence for the prediction error-based learning in crickets *Gryllus bimaculatus*

（コオロギを用いた予測誤差に基づく学習の検証）

Kanta TERAO

（寺尾 勘太）

March 2018

Contents

General Introduction

To adapt to an ever-changing environment, animals have evolved capabilities to store information on the environment from their experiences. The fact that various animals have such learning ability[1-8] (mammals, birds, fishes, mollusks, insects and many other animals including planarians) suggests that learning is important for survival of animals. An important research subject is to clarify learning mechanisms and to understand how such capabilities and underlying mechanisms have evolved.

One of the most thoroughly studied forms of learning is Pavlovian (or classical) conditioning. In classical conditioning, first discovered by Pavlov in dogs[9], animals associate a neutral stimulus (conditioned stimulus, CS) with a biologically significant stimulus (unconditioned stimulus, US), the latter of which induces a behavioral response of the animal. Early learning theorists proposed that temporal "contiguity" is critically important to achieve classical conditioning[10]. In this proposal, it is stated that the more closely together in time a CS and a US occur, the more likely will the association between CS and US be formed.

Later, learning theorists suggested that temporal contiguity is not sufficient for classical conditioning. Critical evidence for this was obtained from the finding of blocking by Kamin[11, 12]. He observed in rats that a stimulus X that had been paired previously with a US could block subsequent association of a second stimulus Y to the US when the two stimuli were paired in compound with the same US (XY+ training). This finding indicates that temporal contiguity between CS and US is not the critical determinant of classical conditioning, because contiguity between Y and US is maintained in XY+ training but does not cause an association between Y and US. Kamin[12, 13] argued that no learning of stimulus Y occurs when the US was fully predicted by stimulus X and argued that "surprise" is needed for learning. This proposition was summarized as the "prediction error" in several learning theories, where λ represents actual US and V represents predicted US, and thus λ -V represents the US prediction error. One of the most

famous theories was proposed by Rescorla and Wagner[14] (Table 1a).

During the time that learning theorists were refining the criteria to achieve classical conditioning based on behavioral experiments, neuroscientists began to study neural mechanisms of learning. Studies in invertebrates have greatly contributed to the clarification of cellular and molecular mechanisms of associative learning since their relatively simple nervous systems greatly facilitated detailed electrophysiological and molecular analyses. Eric Kandel and his colleagues studied neural mechanisms for association of CS and US signals in classical conditioning of gill withdrawal reflex of the mollusc *Aplysia*[5]; they showed that tactile CS signals and electric shock aversive US signals conveyed by a biogenic amine serotonin (5-HT) converge onto a presynaptic site of a specific motor neuron.

Recent studies have suggested similarity of the roles of biogenic amines in mediating US signals in associative learning in different animal taxa. Researchers studying associative learning in honeybees and crickets showed that appetitive and aversive US signals are mediated by a biogenic octopamine (OA) and dopamine (DA), respectively[15-19]. Gene knockout experiments using the CRISPR/Cas9 system and gene knockdown experiments using RNA interference demonstrated that the type 1 octopamine receptor (OA1) and type 1 dopamine receptor (DopR1) mediate appetitive and aversive learning, respectively, in crickets[20, 21]. In the fruit-fly *Drosophila melanogaster*, however, recent studies have suggested that different sets of dopaminergic neurons mediate both appetitive and aversive US signals via the type 1 dopamine receptor DopR1[22-25]. Researchers studying associative learning in rodents and primates showed that reward signals are mediated by DA[26].

Both associative learning theoretical studies and neuroscience studies have led to remarkable successes, but these studies have resulted in different views of classical conditioning by neuroscientists working on *Aplysia* and psychologists working on mammals: Neuroscientists working on *Aplysia* came to a consensus view that detection

of temporal contiguity, or coincidence detection, is sufficient to achieve classical conditioning, in contrast to the widespread view that contiguity is insufficient to achieve associative learning in mammals. Later, neuroscientists working on mammals discovered a possible neural basis of learning by "surprise"; DA neurons in the midbrain appear to mediate signals about reward prediction error during associative learning[1, 27].

Many questions, however, about the basic mechanisms of classical conditioning have remained to be answered. One of the important questions is validity of the prediction error-based learning theory, especially the Rescorla-Wagner model[14]: unambiguous demonstration of the models has not been achieved in any learning systems. Blocking can also be accounted for by theories other than the Rescorla-Wagner model[28-30], and thus experiments are needed to discriminate among different theories. The most influential theory is the attentional theory (or theory of attention) proposed by Machintosh[28] and by Pearce and Hall[29], which is a prediction error-based learning theory and accounts for blocking by loss of attention to a stimulus (Table 1b, c). Another notable theory is the comparator hypothesis[30], which accounts for blocking by cue competition during memory retrieval. Experiments have been performed to discriminate the Rescorla-Wagner model from other theories in some learning systems, but unequivocal evidence to distinguish all alternative theories has not been obtained in any learning systems[31, 32].

Another question is whether DA neurons convey prediction error signals in aversive learning. In mammals, some researchers have suggested that separate classes of midbrain neurons mediate prediction error signals about reward and aversive US, respectively[33, 34], the former being DA neurons but the latter possibly not being DA-ergic. Other researchers, on the other hand, have proposed that a single class of DA neurons integrates reward and aversive US signals to encode value prediction error signals[35]. To my knowledge, a neural basis of learning depending on the prediction error in aversive learning has not been reported in any animals other than vertebrates. Therefore, identification and comparison of neurotransmitters mediating reward and aversiveness remain critical issues in

mammals and in other animals.

In my doctoral thesis, I attempt to provide insights for general learning mechanisms beyond animal taxa. The thesis consists of two chapters. In chapter 1, I show results of a blocking experiment in appetitive learning in crickets. I demonstrated one-trial blocking, which verifies the applicability of the Rescorla-Wagner model and comparator hypothesis. I propose a new learning model to account for cricket learning by modifying a previous model[6]. I evaluated the model by a pharmacological experiment that I call "auto-blocking", which verifies the applicability of prediction error-based learning theories including the Rescorla-Wagner model. Thus, these results provide evidence for the first time that is specifically in favor of the Rescorla-Wagner model but not the other alternative theories. In chapter 2, I show results of a blocking experiment in aversive learning in crickets. I revised the learning model to account for both appetitive and aversive learning and evaluated it by an auto-blocking experiment. The results provided rigorous evidence of prediction error-based learning and suggest that OA and DA convey reward and aversive US prediction error signals, respectively, in crickets.

References

1. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science*, **275**, 1593–1599 (1997).
2. Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87-115 (2006).
3. Wen, C., Ogura, Y. & Matsushima, T. Striatal and tegmental neurons code critical signals for temporal-difference learning of state value in domestic chicks. *Front. Neurosci.* **10**, 1–25 (2016).
4. Aoki, T. et al. Imaging of neural ensemble for the retrieval of a learned behavioral program. *Neuron* **78**, 881–894 (2013).
5. Kandel, E. R. The molecular biology of memory storage: A dialogue between gene and synapses. *Science*, **294**, 1030–1038 (2001).
6. Mizunami, M., Unoki, S., Mori, Y. & Hirashima, D. Roles of octopaminergic and dopaminergic neurons in appetitive and aversive memory recall in an insect. *BMC Biology*, **16**, 1–16 (2009).
7. Prados, J. et al. Cue competition effects in the planarian. *Anim. Cogn.* **16**, 177–186 (2013).
8. Perry, C. J., Barron, A. B. & Cheng, K. Invertebrate learning and cognition: Relating phenomena to neural substrate. *Wiley Interdiscip. Rev. Cogn. Sci.* **4**, 561–582 (2013).
9. Pavlov, I.P, Conditioned reflexes; an investigation of the physiological activity of the cerebral cortex (London: Oxford University Press: Humphrey Milford, 1927).
10. Mazur, J. E. [Chapter 1: History, Background, and Basic Concepts] Learning and behavior [1–25] (Pearson education, Boston, 2013)
11. Kamin, L. J. [" Attention-like" processes in classical conditioning.] [M.R. Jones (Ed.).] [9-31] (Miami Symposium on the Prediction of Behavior, 1967: Aversive Stimulation. Coral Gables, Florida: University of Miami Press 1968)
12. Kamin, L. [Predictability, surprise, attention and conditioning] *Punishment and*

aversive behavior [Campbell, B.A. & Church, R.M. (eds.)] [279-298] (Appleton-Century-Crofts, New York, 1969).

13. Kamin, L.J. [Selective association and conditioning.] *Fundamental issues in associative learning.* [N.J. Mackintosh and W.K. Honig (Eds.)] [42-64] (Proceedings of a symposium held at Dalhousie University, Halifax, June 1968. Halifax: Dalhousie University Press, 1969)
14. Rescorla, R. A., Wagner, A. R. & others. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. Class. Cond. II Curr. Res. theory **2**, 64–99 (1972).
15. Hammer, M. & Menzel, R. Multiple sites of associative odor learning as revealed by local brain microinjections of octopamine in honeybees. *Learn. Mem.* **5**, 146–56 (1998).
16. Unoki, S., Matsumoto, Y. & Mizunami, M. Roles of octopaminergic and dopaminergic neurons in mediating reward and punishment signals in insect visual learning. *Eur. J. Neurosci.* **24**, 2031–2038 (2006).
17. Unoki, S., Matsumoto, Y. & Mizunami, M. Participation of octopaminergic reward system and dopaminergic punishment system in insect olfactory learning revealed by pharmacological study. *Eur. J. Neurosci.* **22**, 1409–1416 (2005).
18. Nakatani, Y. et al. Why the carrot is more effective than the stick: Different dynamics of punishment memory and reward memory and its possible biological basis. *Neurobiol. Learn. Mem.* **92**, 370–380 (2009).
19. Vergoz, V., Roussel, E., Sandoz, J. C. & Giurfa, M. Aversive learning in honeybees revealed by the olfactory conditioning of the sting extension reflex. *PLoS One* **2**, (2007).
20. Awata, H. et al. Knockout crickets for the study of learning and memory: Dopamine receptor Dop1 mediates aversive but not appetitive reinforcement in crickets. *Sci. Rep.* **5**, 15885 (2015).

21. Awata, H. et al. Roles of OA1 octopamine receptor and Dop1 dopamine receptor in mediating appetitive and aversive reinforcement revealed by RNAi studies. *Sci. Rep.* **6**, 29696 (2016).
22. Waddell, S. Neural plasticity: Dopamine tunes the mushroom body output network. *Curr. Biol*. **26**, R102–R124 (2016).
23. Kim, Y.-C., Lee, H.-G. & Han, K.-A. D1 dopamine receptor dDA1 is required in the mushroom body neurons for aversive and appetitive learning in *Drosophila*. *J. Neurosci.* **27**, 7640–7647 (2007).
24. Burke, C. J. et al. Layered reward signaling through octopamine and dopamine in *Drosophila*. *Nature* **492**, 433–437 (2012).
25. Liu, C. et al. A subset of dopamine neurons signals reward for odour memory in *Drosophila*. *Nature* **488**, 512–516 (2012).
26. Wise, R. & Rompre, P. Brain Dopamine and Reward. *Annu. Rev. Psychol.* **40**, 191–225 (1989).
27. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43–48 (2001).
28. Mackintosh, N. J. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol. Rev.* **82**, 276–298 (1975).
29. Pearce, J. M. & Hall, G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532-552 (1980).
30. Miller, R.R. & Matzel L.D. The comparator hypothesis: a response rule for the expression of associations. *Psychol. Learn. Motiv.* **22**, 51-92. (1988)
31. Pearce, J. M. [Associative learning] Animal Learning & Cognition [35–91] (Psychology press, New York, 2008)
32. Miller, R. R., Barnet, R. C. & Grahame, N. J. Assessment of the Rescorla-Wagner model. *Psychol. Bull.* **117**, 363-386 (1995).

33. Fiorillo, C. D. Two Dimensions of Value: Dopamine Neurons Represent Reward But Not Aversiveness. *Science* **341**, 546–549 (2013).
34. Stauffer, W. R., Lak, A., Kobayashi, S. & Schultz, W. Components and characteristics of the dopamine reward utility signal. *J. Comp. Neurol.* **524**, 1699–1711 (2016).
35. Matsumoto, H., Tian, J., Uchida, N. & Watabe-Uchida, M. Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *Elife* **5**, 1–24 (2016).

Tables

Table 1. Learning theories to account for blocking.

Theory	Equation
a. Rescorla-Wagner model[14]	$\Delta V = \alpha (\lambda - V_{\Sigma})$ $V_{\Sigma} = V_A + V_B + \ldots V_X$
b. Attentional theory by Mackintosh[28]	$\Delta V = \alpha_A (\lambda - V_A)$ α_A is positive if $\lvert \lambda - V_A \rvert < \lvert \lambda - V_X \rvert$ α_A is negative if $\lvert \lambda - V_A \rvert \geqq \lvert \lambda - V_X \rvert$
c. Attentional theory by Pearce and Hall[29]	$\Delta V_A = S_A \alpha_A \lambda$ $\alpha_A^n = \lvert \lambda^{n-1} - V_{\Sigma}^{n-1} \rvert$

V: associative strength, which corresponds to US prediction; ΔV: change in associative strength in a particular trial; α: learning-rate parameter; λ: magnitude of the US; S: intensity of the stimulus; Description of equations follows Pearce and Hall[29].

Chapter I

Critical evidence for the validity of the Rescorla-Wagner model in associative learning

Note: This chapter includes materials modified from

Terao, K., Matsumoto, Y. & Mizunami, M. Critical evidence for the prediction error theory in associative learning. Sci. Rep. 5, 8929 (2015),

Abstract

In associative learning in mammals, it is widely accepted that the discrepancy, or error, between actual and predicted reward determines whether learning occurs. One of the most influential prediction error-based learning theory is Rescorla and Wagner model. Complete evidence for the Rescorla and Wagner model, however, has not been obtained in any learning systems. Rescorla and Wagner model stems from the finding of a blocking phenomenon, but blocking can also be accounted for by other theories, such as the attentional theory and comparator hypothesis. I demonstrated blocking in classical conditioning in crickets and obtained evidence to reject the attentional theory. To obtain further evidence verifying the Rescorla and Wagner model, I constructed a neural model to match the prediction error-based learning theories, by modifying a previous model of learning in crickets, and I tested a prediction from the model. The model predicts that pharmacological intervention of octopaminergic transmission during appetitive conditioning impairs learning but not formation of reward prediction itself, and it thus predicts no learning in subsequent training. I observed such an "auto-blocking", which could be accounted for by the Rescorla and Wagner model but not by comparator hypothesis. This study unambiguously demonstrates validity of the Rescorla and Wagner model in associative learning.

Introduction

Understanding computational rules underlying associative learning, such as classical conditioning and operant conditioning, is a major goal of neuroscience. In associative learning in mammals, it is widely accepted that the discrepancy, or error, between the actual unconditioned stimulus (US) and the predicted US determines whether learning occurs when a stimulus is paired with the US[1]. This theory stems from the finding of "blocking" by Kamin[2, 39]. He observed, in rats, that a stimulus X that had been paired previously with a US could block subsequent association of a second stimulus Y to the US when the two stimuli were paired in compound with the same US. Kamin[2, 40] argued that the blocking is due to the requirement of surprise for learning, i.e., no learning occurs when the US is fully predicted, and this proposition was formulated into the prediction error-based learning theories. One of the most influential prediction error-based learning theory is proposed by Rescorla and Wagner[3] (see General Introduction). Recent neuroscience research in mammals has demonstrated that activities of dopamine (DA) neurons in the ventral tegmental area of the midbrain mediate prediction error signals in classical conditioning[1, 4] and instrumental conditioning[5].

Unambiguous demonstration of the Rescorla-Wagner model, however, has not been achieved in any learning systems. Blocking can also be accounted for by theories other than the Rescorla-Wagner model[6-8], and thus experiments are needed to discriminate among different theories. One of the important alternative theories is the attentional theory (or theory of attention) proposed by Machintosh[9] and Pearce and Hall[10], which is prediction error-based learning model and accounts for blocking by a loss of attention to a stimulus in the training trial. Another notable theory is the comparator hypothesis[11], which accounts for blocking by cue competition during memory retrieval. Experiments have been performed to discriminate the Rescorla-Wagner model from other theories in some learning systems[6-8], but unequivocal evidence to reject all alternative theories has not been obtained in any learning systems.

Studies in invertebrates have greatly contributed to an understanding of cellular and

molecular mechanisms of associative learning[12-16], but whether the prediction error-based learning is achieved in invertebrates has not been examined. One of the reasons for the lack of such study is the difficulty in establishing experimental procedures to convincingly demonstrate blocking. In insects, for example, earlier studies in honey bees showed a blocking-like effect in free-flight stimulus selection experiments[17,18] and in classical conditioning of proboscis extension responses[19-21], but subsequent studies have failed to establish blocking as a robust learning phenomenon[22-24]. Another reason is that, even though blocking has been established in some invertebrates, especially mollusks[25-28], experiments have not been performed to discriminate different theories of blocking in any invertebrate species.

In this study, I obtained unequivocal evidence of blocking in classical conditioning in crickets. Crickets are newly emerging experimental animals, in which neural mechanisms of classical conditioning have been unraveled in some detail[29]. For example, I observed that octopamine (OA) receptor antagonists impair appetitive learning but not aversive learning, whereas DA receptor antagonists impair aversive learning but not appetitive learning[30,31]. Moreover, I observed that OA receptor antagonists impair retrieval of appetitive memory but not that of aversive memory, whereas DA receptor antagonists impair retrieval of aversive memory but not that of appetitive memory[32]. Therefore, I concluded that OA neurons and DA neurons code appetitive and aversive US in conditioning and that activation of these neurons is also needed for retrieval of appetitive and aversive memory, respectively, and proposed a neural model of classical conditioning to account for these findings[32]. In this study, I performed behavioral analysis of blocking in crickets. Moreover, I proposed a model that matches the prediction error-based learning by modifying a previous model and performed pharmacological tests of a prediction from the model. My results provide rigorous evidence for the prediction error-based learning in invertebrates, and verified the Rescorla-Wagner model by rejecting alternative theories for the first time.

Results

Effects of compound conditioning. In experiments to study blocking, I used odor-pattern compound conditioning (OP+ conditioning), in which a compound stimulus consisting of an odor (O) and a visual pattern (P) is paired with a water US (reward) (+) (Table 1). The procedures for experiments are described in the Methods section and illustrated as parts of Fig. 1. As a prerequisite for such experiments, I first tested whether OP+ compound training leads to learning of both the odor and the visual pattern. One group of animals was subjected to 4-trial OP+ compound training and another group (control group) was subjected to 4-trial olfactory conditioning (O+ conditioning). In both groups, the relative preference for the conditioned odor and control odor was tested before and at 30 min after training. I used a generalized linear mixed model (GLMM) to evaluate the data (see Methods). The results are shown in Fig. 2a. Both the compound group and control group exhibited significantly increased preference for the conditioned odor after training compared to that before training (test term, $p = 2 * 10^{-16}$, $z = 12.46$, see Supplemental table S1). The preference for the conditioned odor after training in the compound group did not significantly differ from that in the control group (test * training term, $p = 0.638$, $z = -0.471$). The results showed that learning was achieved in the compound group as in the control group. I thus conclude that odor-pattern compound conditioning leads to conditioning of the odor.

Next, one group of animals was subjected to 4-trial OP+ compound training and another group (control group) was subjected to 4-trial visual pattern conditioning (P+ training). The results are shown in Fig. 2b. Both the compound group and control group exhibited significantly increased preference for the conditioned pattern after training compared to that before training (test term, $p = 2 *10^{-16}$, $z = 9.867$). Thus, odor-pattern conditioning also leads to conditioning of the pattern. In addition, I observed that the preference for the conditioned pattern after training in the compound group was significantly higher than that in the control group (test * training term, $p = 2.29 *10^{-3}$, $z = -3.050$). I will discuss for the results in discussion section of

the Chapter 2.

Demonstration of blocking. I next studied whether blocking occurs in crickets. One group of animals (blocking group) was subjected to 4-trial P+ training and then 4-trial OP+ training (Table 1). Another group (unpaired control group) was subjected to unpaired presentations of a visual pattern and reward (P / +) 4 times each and then 4-trial OP+ training. The results are shown in Fig. 3a. The preference for the trained odor after training in the control group was significantly higher than that before training in the same group and was also significantly higher than that before or after training in the blocking group (test * training term, $p = 3.73 * 10^{-13}$, $z = 7.265$). The results showed blocking effect that conditioning was achieved in the control group but not in the blocking group. The results demonstrate blocking of olfactory learning..

Similarly, I observed blocking of visual pattern learning. One group of animals was subjected to 1-trial O+ training and then 4-trial OP+ training. In this experiment, 1-trial O+ training was sufficient because O+ training is more effective than P+ training[30]. Another group was subjected to unpaired presentations of an odor and reward (O / +) once and then 4-trial OP+ training. The results are shown in Fig. 3b. Both the blocking group and unpaired control group exhibited significantly increased preference for the conditioned pattern after training compared to that before training (test term, $p = 0.0371$, $z = 2.085$). In addition, I observed that the preference for the conditioned pattern after training in the unpaired control group was significantly higher than that in the blocking group (test * training term, $p = 1.1 *10^{-5}$, $z = 4.396$). The results demonstrate blocking of visual learning.

Evaluation of the attentional theory: demonstration of blocking with 1-trial compound conditioning. Rescorla-Wagner model is a predominant theory to account for blocking[1], but a few other theories can also explain blocking. Another influential one is attentional theory, which accounts for blocking by a loss of attention to a stimulus[9,10]. I investigated which theory, the

attentional theory or the Rescorla-Wagner model, better accounts for blocking in crickets; other theories will be discussed later. The decisive method to discriminate attentional theories from alternative theories is to study the effect of blocking with X+ training and subsequent 1-trial XY+ training: the attentional theory predicts a loss of attention to Y after the first XY+ training (not after X+ training), and thus 2-trial XY+ training is needed for achieving blocking[27]. In contrast, alternative theories including Rescorla-Wagner model predicts blocking with 1-trial XY training. The results are shown in Fig. 4a. The preference for the trained odor after training in the compound and unpaired control group was significantly higher than that before training in the same group and was also significantly higher than that before or after training in the blocking group (test * training term, compound: $p = 4.09 * 10^{-8}$, $z = 5.487$; unpaired: $p = 4.19 * 10^{-7}$, $z = 5.060$). The results showed blocking effect that conditioning was achieved in the compound and unpaired control group but not in the blocking group. The occurrence of blocking with 1-trial OP+ training better match alternative theories other than the attentional theory.

Alternatively, blocking with X+ training and subsequent 1-trial XY+ training might be due to a loss of attention to Y, which is presented together with X. This "naïve" attentional theory has not been seriously considered in mammals because it is thought to be too simplistic[10], but this theory deserves consideration for blocking in crickets. In order to evaluate whether crickets' attention to an odor is lost after P+ training, their behavioral responses to an odor were observed. Crickets often extend and swing their maxillary palpi vigorously when they have perceived a food odor, which I refer to as maxillary palpi extension response (MER), but they do not exhibit MER when a visual pattern is presented. I compared the percentage of MER (%MER) to OP compound and that to a pattern alone to estimate crickets' response to the odor in the OP compound after P+ training. Two groups of animals were subjected to 4-trial P+ training and then, in one group, MER was tested to P alone and then to OP compound. In another group, the sequence of tests was reversed. Because the sequence of the tests had no effect on responses to

OP compound or to P alone, the data from the two groups were pooled as a paired presentation group. The paired group exhibited a very low %MER to P alone but a high %MER to OP compound ($\chi^2 = 9.3889$, $p = 0.0022$, McNemar's test; Fig. 4b), suggesting that MERs to OP compound are, in most part, caused by the odor. Another group (unpaired control group) was subjected to unpaired presentation of a pattern and reward (P / +) 4 times each, and then MERs to OP compound were tested. This group exhibited a high %MER to OP compound. The %MER to OP compound did not significantly differ between the P+ group and unpaired group ($p = 0.43$, Fisher's exact test). Assuming that MERs to OP compound are in large part due to the odor, this observation suggests that P+ training does not attenuate attention to the odor in OP compound. Thus, these observations rejected the naïve attentional theory.

A neural circuit model of classical conditioning that matches prediction error-based learning theories. The results described above support the Rescorla-Wagner model but do not unequivocally prove it, since the attentional theory is not the only alternative theory to account for blocking[6-8]. In order to obtain further evidence supporting the Rescorla-Wagner model and to discriminate it from alternative theories, I constructed a model of classical conditioning that matches the prediction error-based learning (Supplementary Fig. S1a) by modifying a previous model of classical conditioning in crickets[32] (Supplementary Fig. S1b). An essential assumption in my new model is that enhancement of synaptic transmission from "CS" neurons to three classes of neurons, i.e., "CR", "OA1" and "OA2" neurons, are necessary for achieving appetitive conditioning, in which "CS" neurons are neurons mediating CS, "CR" neurons are neurons whose activation leads to CR and "OA1" and "OA2" neurons are different classes of octopaminergic (OA) neurons mediating appetitive US. Other assumptions in the model and how this model matches the prediction error-based learning theories are described in legends of Supplementary Fig. S1. To better account for the model, information coded by "OA1" and "OA2" neurons is shown in Supplementary Table S2. I used this model for designing an

experiment to obtain further evidence of Rescorla-Wagner model.

Demonstration of auto-blocking. I noticed that my model predicts that blockade of synaptic transmission from OA neurons by an OA receptor antagonist during a pairing of a stimulus (Y) with reward (Y+ training) impairs learning of Y but not formation of reward prediction by Y. This is because it impairs enhancement of "CS-CR" synapses but not that of "CS-OA1" and "CS-OA2" synapses in the model, and enhancement of all of these synapses are necessary for appetitive learning but that of "CS-OA1" synapses is sufficient for formation of reward prediction (see Supplementary Table S2 and legends of Supplementary Fig. S1). Therefore, subsequent Y+ training, after recovery from the synaptic blockade caused by epinastine, should produce no learning. This effect can be termed "auto-blocking", because learning of Y is blocked by US prediction by Y itself, not by X in the case of blocking.

I tested whether auto-blocking occurs in crickets. I used epinastine, an antagonist of the insect OA receptor[33], which impairs appetitive learning but not aversive learning in crickets[30,31]. One group of animals (auto-blocking group) was injected with epinastine into the head haemolymph and the group was subjected to 4-trial O+ training 30 min later. The timing of injection and the concentration of the drug were based on a previous study[30]. The next day, the group was subjected to 1-trial O+ training. Another group (unpaired control group) was subjected to unpaired presentation of the odor and reward (O / +) 4 times each with application of epinastine and was subjected to 1-trial O+ training the next day. The results are shown in Fig. 5. Both the auto-blocking group and unpaired control group exhibited significantly increased preference for the conditioned odor after training compared to that before training (test term, $p = 3.25 * 10^{-4}$, $z = 3.594$). In addition, I observed that the preference for the conditioned odor after training in the unpaired control group was significantly higher than that in the auto-blocking group (test * training term, $p = 2.23 *10^{-6}$, $z = 4.732$). The results demonstrate auto-blocking of olfactory learning. The results demonstrate that auto-blocking occurs in crickets,

providing evidence for the prediction error-based learning in classical conditioning in crickets. Non-prediction error-based alternative hypotheses, including comparator hypothesis and “naive” attentional theory, do not match auto-blocking. Demonstration of blocking and auto-blocking provide critical evidence for the validity of the Rescorla-Wagner model.

Discussion

I obtained convincing evidence for the Rescorla-Wagner model. At first, I obtained evidence of blocking, i.e., no learning of a stimulus (Y) by pairing of a compound of Y and another stimulus (X) with reward (XY+ training) when it is preceded by X+ training, in crickets. Among theories to account for blocking, I focused on Rescorla-Wagner model[3] and attentional theories[9,10], the former accounting for blocking by lack of US prediction error and the latter by lack of attention to Y. The results of my experiment with 1-trial XY+ conditioning support the former theory but not the latter. In order to obtain further evidence for the Rescorla-Wagner model, I constructed a neural circuitry model of classical conditioning that is consistent with the prediction error-based learning theories (Supplementary Fig. S1a), by revising a previous model[32] (Supplementary Fig. S1a). How this model matches the prediction error-based learning theories is shown in Supplementary Table S2 and how it accounts for blocking is illustrated in Fig. 6. The new model predicts that application of an OA receptor antagonist before Y+ training impairs learning of Y but not formation of reward prediction by Y. In accordance with this prediction, I observed no learning of Y by subsequent Y+ training. The finding of "auto-blocking" in crickets can be accounted for by the prediction error-based learning theories but not by other competitive theories to account for blocking including comparator theory (see below). Evidence of blocking and auto-blocking provide rigorous evidence for validity of the Rescorla-Wagner model.

My results provide the first robust evidence for prediction error-based learning in invertebrates. In mammals, a wide range of brain regions is implicated for calculating US prediction error relying on the prediction for future US[41], thus, one might argue that such computation is formidable for the small brains of insects. It should be noted, however, that my model suggests that computation of US prediction error can be achieved by a simple neural circuit consisting of a small number of elements.

It has remained controversial whether blocking occurs in learning of insects. Earlier studies

using free-flight honey bees showed a blocking-like effect[17,18], but in recent studies, the effect has been concluded to be due to confounding factors[23,24]. In olfactory conditioning of the proboscis extension response in harnessed honey bees, earlier studies also showed a blocking-like effect[19-21], but in recent comprehensive studies, it was concluded that blocking is not a robust phenomenon[22]. Attempts to demonstrate blocking failed in the fruit-fly *Drosophila*[34,35]. These negative reports, however, do not necessarily indicate that blocking does not occur in these insects. Rather, they may indicate that more effort is needed to establish experimental paradigms to demonstrate blocking.

Suitable controls are needed to discriminate blocking from confounding factors, and there has been debate concerning the most appropriate control procedure to demonstrate blocking[23,28]. Considering the debate, I performed three different comparisons to demonstrate blocking, namely, I showed that (1) the blocking group exhibited no learning by comparing preferences for the CS before and after training, (2) preference for the CS after training in the blocking group was significantly less than that in the unpaired group, and (3) it was significantly less than that in the compound group. In the original experiment on blocking, Kamin[2] used a group with XY+ compound conditioning as the control group, and this is still considered a typical control procedure. Some researchers, on the other hand, prefer to use unpaired presentation of X and US (+) before XY+ training in order to equalize the amount of exposure to X and US between the control group and the blocking group[16]. Other researchers argue that the use of between-group comparison is problematic and prefer within-group comparison[28]. Regardless of such debate, my study unequivocally demonstrates blocking in crickets.

My model predicts “auto-blocking”, and I indeed observed this phenomenon. Importantly, my auto-blocking experiment demonstrated prediction error-based learning without using XY+ training and thus without stimulus competition between X and Y. To my knowledge, naïve attentional theory and comparator hypothesis[11] assume stimulus competition to account for blocking, and thus they fail to account for auto-blocking.

I propose that the auto-blocking procedure is applicable to learning systems of animals other than crickets. I noticed that auto-blocking can be predicted from computational models other than mine, such as a model proposed by Goel and Gelperin[36]. In the model, activity of the neurons mediating reinforcement signals (they call FN neurons) is like that of US prediction error-based learning theories predict in the blocking experiment. Their model differs in many respects from my model and thus does not account for learning in crickets, but it predicts that blockade of synaptic output from reinforcing neurons impairs learning but not formation of US prediction error-like signals. Moreover, an auto-blocking experiment can be performed in any learning systems if the neurotransmitter of reinforcing neurons is known and synaptic output from reinforcing neurons can be blocked during learning. A combination of blocking and auto-blocking experiments may become a useful procedure for demonstration of the prediction error-based learning.

Although my model shown in Supplementary Fig. S1a focused on roles of OA neurons in conveying prediction error signals for appetitive US, I can similarly assume roles of dopamine (DA) neurons in conveying prediction error signals for aversive US, and confirmation of these is one of my major goals. I hypothesize that OA and DA neurons projecting to the lobes of the mushroom body convey prediction error signals for olfactory conditioning, because the lobes have been suggested to be the sites of association between CS and US[14,37]. It would be interesting to compare activities of these neurons in insects to those of midbrain DA neurons in mammals, which have been suggested to convey prediction error signals for classical conditioning[1] and instrumental conditioning[5].

Neural circuitry mechanisms for computation of the prediction error has been studiously investigated, especially in mammals[1, 41]. Crickets should emerge as new pertinent model animal to elucidate this important subject because of relatively simple and small nervous system. My model predicts that (1) there should be two types of OA or DA neurons, one type being inhibited and the other type being excited by CS presentation after conditioning and (2) the former

conveys US prediction error and activation of them is needed for enhancement of one type of synapses, whereas activation of the latter is needed for memory retrieval. These predictions should guide future electrophysiological studies. It should also be noted that my model indicates that the prediction error-based learning theories does not account for all aspects of associative learning in crickets. The model assumes synaptic plasticity in three different synapses in the circuitry and suggests that the plasticity of one type of synapses ("CS-CR" synapses) is governed by US prediction error but that of other synapses ("CS-OA1" synapses and "CS-OA2" synapses) is governed by US (see Supplementary Fig. S1). Therefore, actual mechanisms of associative learning in insects may be more elaborate than the prediction error-based learning theories assumes. This argument may be applicable to mammals; because the prediction error-based learning theories is known to account for some but not all features of associative learning in mammals[6-8], further effort is needed to delineate the extent by which the prediction error-based learning theories accounts for mechanisms of associative learning in mammals.

Methods

Insects. Adult male crickets, *Gryllus bimaculatus*, at 1 week after the imaginal molt were used. Before the experiment, animals were placed individually in beakers and deprived of drinking water for 4 days to enhance their motivation to search water.

Olfactory and Visual Conditioning Procedure. I used classical conditioning and operant testing procedures described previously (Fig. 1)[30,31]. In olfactory conditioning, peppermint odor was used as conditioned stimulus (CS), while in visual conditioning, a white-center and black-surround pattern (white-center pattern) was used as CS. In compound conditioning, the odor and pattern were presented simultaneously (compound CS). Water was used as US (reward). A syringe containing water was used to present CS and US to each cricket. A filter paper soaked with a drop of peppermint essence or a white-center pattern was attached to the needle (code number NN-2238S, TERUMO, Tokyo) of the syringe (Fig. 1). For pairing trial, an odor and/or visual pattern was approached to the antennae or the head and held for 3 sec, and subsequently a drop of water was attached to its mouth. Crickets typically received 4-trial training with an inter-trial interval (ITI) of 5 min. After olfactory or compound conditioning trials, the air in the beaker was ventilated.

Preference Tests. The odor and pattern preference tests were carried out as described elsewhere[30,31]. All groups were tested with relative preference between the conditioned (peppermint) and control (vanilla) odor or between the conditioned (white-center) and control (black-center) pattern from 1 day to 1 hour before and at 30 min after conditioning. In the chamber to test odor preference, the floor had two holes that connected the chamber with two cylindrical odor sources containing a filter paper soaked with either a drop of peppermint (Mikoya, Tokyo) or vanilla essence (Kyoritsu-foods, Tokyo) diluted with 5 drops of water and covered with fine gauze net (Fig. 1). Three containers were mounted on a rotative holder and

two of the three sources could be located simultaneously beneath the holes of the test chamber. In the apparatus for the pattern preference test, two white-center patterns and one black-center pattern were displayed on a gray sliding wall at the end of the test chamber, and a white-center pattern and a black-center pattern could be presented at the same time (Fig. 1). Before testing, a cricket was transferred to the waiting chamber and left for about 4 min to become accustomed to the surroundings. Then the cricket was allowed to enter the test chamber and then test started. Two min after the test start, the relative positions of the odor sources or patterns were changed by either rotating the container holder or sliding the wall. The preference test lasted for 4 min. I considered that the cricket visited an odor source when the cricket probed the top net with its mouth or palpi. Previous preliminary experiments suggested that the observer effects are minimal; evaluation of visiting time of crickets in the test by different experimenters yielded almost identical results, with differences of <5% (Matsumoto, personal communication, 2017). If the total visiting time of a cricket to odor sources or patterns was less than 10 s, I considered that the animal was less motivated, possibly due to a poor physical condition, and the data were rejected. Animals that fell into this category comprised less than 10% of the total.

I measured maxillary palpi extension response (MER) to observe crickets' attention to an odor in the odor-pattern compound (see Fig. 1). Crickets often extend and shake their maxillary palpi vigorously when a small filter paper soaked with an essence of a food-related odor is approached to their antennae or when water or sucrose solution has been attached to the mouth or antennae. I recorded MER if a cricket extended its maxillary palpi during 3-sec period in which a small filter paper soaked with odor essence was presented within 1 cm of the antennae.

Pharmacology. Animals were injected with 3 μl of saline containing 2 μM epinastine (Sigma-Aldrich, Tokyo) into the head hemolymph. The estimated final concentration after circulation is 7.0 nM[30].

Statistical Analysis. In the arena test, I considered that an odor or pattern was visited when the cricket probed it with its mouth or palpi. The time visiting each odor or pattern was measured cumulatively. In previous studies, non-parametric statistical tests are used for evaluation of the relative preference. Since it has been proposed that the use of a generalized linear mixed model (GLMM) is advantageous for evaluation of biological data[42], I used a GLMM with a binomial distribution of the relative preference, determined by the search time data sampled for each second, and logit link function. I included the test condition (test before or after training), training procedure and the interaction term (test * training) as fixed effects in the GLMM, with the training and test terms being categorical variables. Individual cricket was used as a random effect, allowing the random intercept. I used R (ver. 3. 3. 1) and lme4 (ver. 1.1.12) packages for statistical analysis. I refer to as "significantly different" if p-values in the Wald statistical test in the GLMM were $p<0.05$.

The percentage of MER was calculated as the number of crickets that showed MER to an odor or pattern with respect to the total number of crickets tested. The difference in response level to the CS was evaluated by means of a McNemar test. Differences in CS responses between groups were assessed using Fisher's exact tests.

References

1. Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87-115 (2006).
2. Kamin, L. [Predictability, surprise, attention and conditioning] *Punishment and aversive behavior* [Campbell, B.A. & Church, R.M. (eds.)] [279-298] (Appleton-Century-Crofts, New York, 1969).
3. Rescorla, R.A. & Wagner, A.R. [A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement] *Classical Conditioning II* [Black, A. & Prokasy, W.R. (eds.)] [64-99] (Academic Press, New York, 1972).
4. Steinberg, E.E., *et al*. A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* **16**, 966-973 (2013).
5. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412,** 43-48 (2001).
6. Miller, R.R., Barnet, R.C., & Grahame, N.J. Assessment of the Rescorla-Wagner model. *Psychol. Bull.* **117**, 363-386 (1995).
7. Pearce, J.M. [Associative learning] *Animal Learning & Cognition* [35-91] (Psychology press, New York, 2008).
8. Mazur J.E. [Chapter 4: Theories and research on classical conditioning] *Learning and behavior* [75-100] (Pearson education, Boston, 2013).
9. Mackintosh, N.J. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol. Rev.* **82,** 276-298 (1975).
10. Pearce, J. M. & Hall, G. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532-552 (1980).
11. Miller, R.R. & Matzel L.D. The comparator hypothesis: a response rule for the expression of associations. *Psychol. Learn. Motiv.* **22**, 51-92. (1988)
12. Heisenberg M. Mushroom body memoir: from maps to models. *Nat. Rev. Neurosci.* **4**, 266-

275 (2003).

13. Menzel, R., Leboulle, G. & Eisenhardt, D. Small brains, bright minds. *Cell* **124**, 237-239 (2006).
14. Davis, R.L. Traces of *Drosophila* memory. *Neuron* **70**, 8-19 (2011).
15. Kandel, E.R. The molecular biology of memory storage: a dialogue between genes and synapses. *Science* **294**,1030-1038 (2001).
16. Crow, T. Pavlovian conditioning of *Hermissenda*: current cellular, molecular, and circuit perspectives. *Learn. Mem.* **11**, 229-38 (2004).
17. Couvillon, P.A. & Arakaki, L. Intramodal blocking in honeybees. *Anim. Learn. Behav.* **25**, 277-282 (1997).
18. Couvillon, P.A., Campos, A.C., Bass, T.D. & Bitterman, M.E. Intermodal blocking in honeybees. *Q. J. Exp. Psychol. B.* **54**, 369-381 (2001).
19. Smith, B.H. & Cobey, S. The olfactory memory of the honeybee *Apis mellifera*. II. Blocking between odorants in binary mixtures. *J. Exp. Biol.* **195**, 91-108 (1994).
20. Thorn, R.S. & Smith, B.H. The olfactory memory of the honeybee *Apis mellifera*. III. Bilateral sensory input is necessary for induction and expression of olfactory blocking. *J. Exp. Biol.* **200**, 2045-2055 (1997).
21. Hosler, J.S. & Smith, B.H. Blocking and the detection of odor components in blends. *J. Exp. Biol.* **203**, 2797-2806 (2000).
22. Guerrieri, F., Lachnit, H., Gerber, B. & Giurfa, M. Olfactory blocking and odorant similarity in the honeybee. *Learn. Mem.* **12**, 86-95 (2005).
23. Blaser, R.E., Couvillon, P.A. & Bitterman, M.E. Blocking and pseudoblocking: new control experiments with honeybees. *Q. J. Exp. Psychol.* **59**, 68-76 (2006).
24. Blaser, R.E., Couvillon, P.A. & Bitterman, M.E. Within-subjects experiments on blocking and facilitation in honeybees (*Apis mellifera*). *J. Comp. Psychol.* **122**, 373-378 (2008).
25. Sahley, C., Rudy, J.W. & Gelperin, A. An analysis of associative learning in a terrestrial

mollusc. *J. Comp. Physiol. A.* **144**, 1-8 (1981).

26. Rogers, R.F. & Matzel, L.D. Higher-order associative processing in *Hermissenda* suggests multiple sites of neuronal modulation. *Learn. Mem.* **2**, 279-298 (1996).
27. Acebes, F., Solar, P., Carnero, S. & Loy, I. Blocking of conditioning of tentacle lowering in the snail (*Helix aspersa*). *Q. J. Exp. Psychol.* **62,** 1315-1327 (2009).
28. Prados, J., *et al*. Blocking in rats, humans and snails using a within-subjects design. *Behav. Process.* **100**, 23-31 (2013).
29. Mizunami, M., Matsumoto, Y., Watanabe, H. & Nishino, H. [Chapter 41: Olfactory and visual learning in cockroaches and crickets] *Invertebrate Learning and Memory,* [Menzel, R. & Benjamin, P. R. (eds.)] [547-558] (Springer, Berlin, 2013).
30. Unoki, S., Matsumoto, Y. & Mizunami, M. Participation of octopaminergic reward system and dopaminergic punishment system in insect olfactory learning revealed by pharmacological study. *Eur. J. Neurosci.* **22**, 1409-1416 (2005).
31. Unoki, S., Matsumoto, Y. & Mizunami, M. Roles of octopaminergic and dopaminergic neurons in mediating reward and punishment signals in insect visual learning. *Eur. J. Neurosci.* **24**, 2031-2038 (2006).
32. Mizunami, M., *et al.* Roles of octopaminergic and dopaminergic neurons in appetitive and aversive memory recall in an insect. *BMC Biol.* **7**, 46 (2009).
33. Roeder, T., Degen, J. & Gewecke, M. Epinastine, a highly specific antagonist of insect neuronal octopamine receptors. *Eur. J. Pharmacol.* **349**, 171–177 (1998).
34. Brembs, B. & Heisenberg, M. Conditioning with compound stimuli in *Drosophila melanogaster* in the flight simulator. *J. Exp. Biol.* **204**, 2849-2859 (2001).
35. Young, J.M., Wessnitzer, J., Armstrong, J.D. & Webb, B. (2011). Elemental and non-elemental olfactory learning in *Drosophila*. *Neurobiol. Learn. Mem.* **96**, 339-352.
36. Goel, P & Gelperin, A. A neuronal network for the logic of *Limax* learning. *J. Comput. Neurosci.* **21**, 259-270 (2006).

37. Liu, C., *et al.* A subset of dopamine neurons signals reward for odour memory in *Drosophila*. *Nature* **488**, 512-516 (2012).

38. Matsumoto, Y. & Mizunami, M. Temporal determinants of long-term retention of olfactory memory in the cricket *Gryllus bimaculatus*. *J. Exp. Biol.* **205**, 1429-1437 (2002).

39. Kamin, L. J. [" Attention-like" processes in classical conditioning.] [M.R. Jones (Ed.).] [9-31] (Miami Symposium on the Prediction of Behavior, 1967: Aversive Stimulation. Coral Gables, Florida: University of Miami Press 1968)

40. Kamin, L.J. [Selective association and conditioning.] *Fundamental issues in associative learning.* [N.J. Mackintosh and W.K. Honig (Eds.)] [42-64] (Proceedings of a symposium held at Dalhousie University, Halifax, June 1968. Halifax: Dalhousie University Press, 1969)

41. Tian, J. *et al.* Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons. *Neuron.* **91,** 1374–1389 (2016).

42. Warton, D. I. & Hui, F. K. C. The arcsine is asinine: the analysis of proportions in ecology. *Ecology* **92**, 3-10 (2011).

Figures and Tables

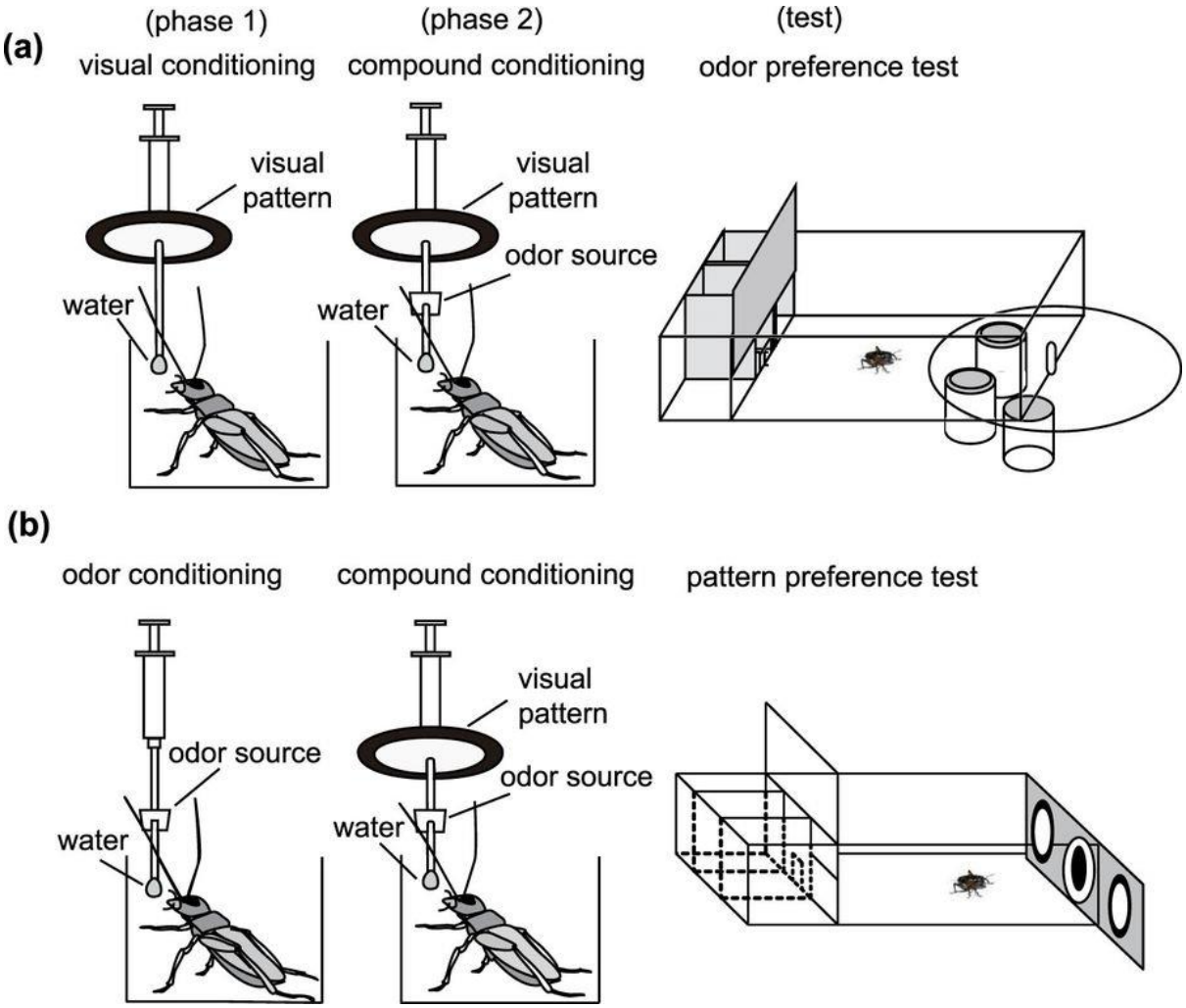


Figure 1. Procedures for blocking experiments. (a) Procedures for the study of blocking of olfactory learning in crickets. A group of water-deprived crickets, individually placed in a beaker, was subjected to pairings of a visual pattern with water US (P+ training) as phase 1 training, and then pairings of an odor-pattern compound with water US (OP+ training) as phase 2 training. Relative preference for the conditioned odor and a control odor was tested before and after training in a test chamber. (b) Procedures for the study of blocking of visual pattern learning. Another group of crickets was subjected to pairings of an odor with water US (O+ training) and then pairings of an odor-pattern compound with water (OP+ training). Relative preference for the conditioned pattern and a control pattern was tested before and after training

in a test chamber.

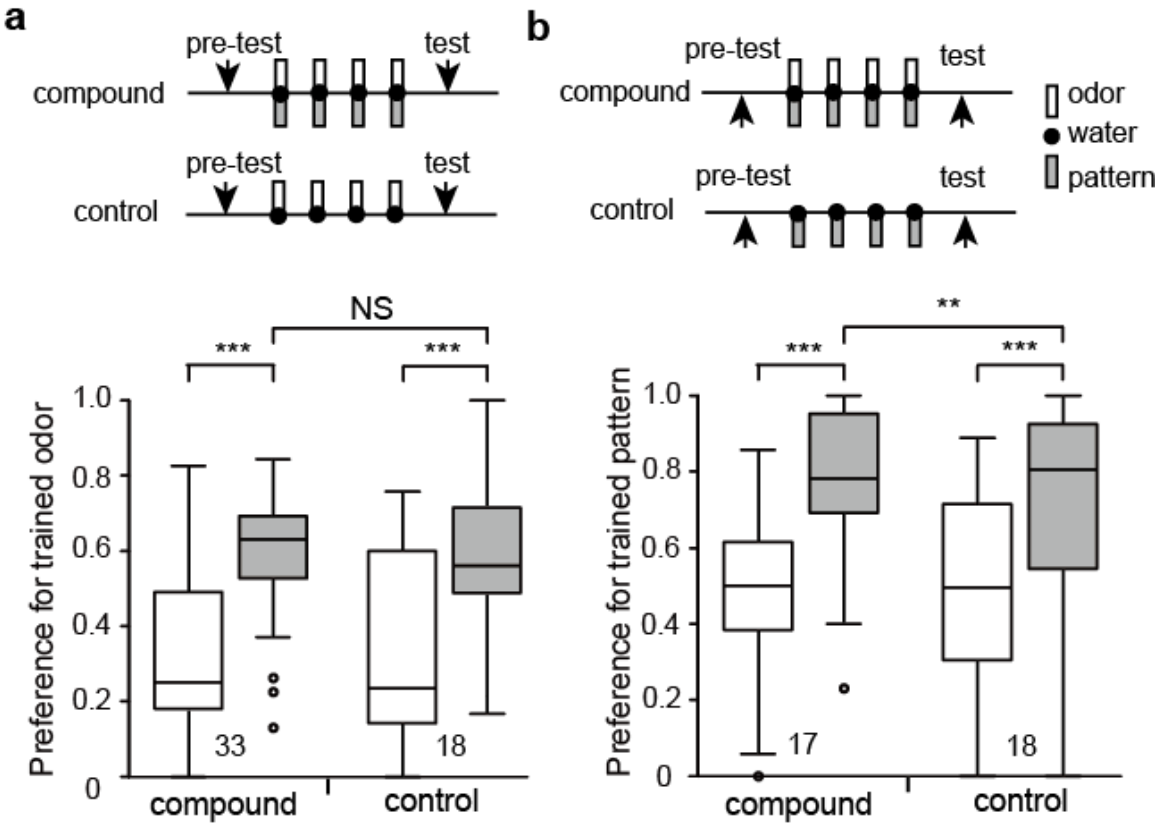


Figure 2. Effect of odor-pattern compound conditioning. (a) Olfactory learning by odor-pattern compound conditioning. One group of animals (compound group) was subjected to 4-trial pairing of an odor-pattern compound with water US and another group (control group) was subjected to 4-trial pairing of an odor alone with water US. (b) Visual learning by odor-pattern compound conditioning. One group was subjected to 4-trial pairing of an odor-pattern compound with water and another group was subjected to 4-trial pairing of a pattern with water. The inter-trial interval (ITI) was 5 min. Relative preference for conditioned odor and control odor or for conditioned pattern and control pattern was tested before and at 30 min after training. The experimental procedures are illustrated at the top, and preference indexes (PIs) for the rewarded odor or pattern before (white bars) and after (grey bars) training are shown as box and whisker diagrams at the bottom. The horizontal line in the box is the median and the box represents the 25-75 percentiles in this and in all following figures. Whiskers extend to extreme values as long as they are within a range of 1.5× box length. The outliers are shown as open

circles. The number of animals is shown below the boxes. A GLMM was used to examine relative preferences for the trained odor or pattern before and after training in the compound and control groups (Supplemental table S1). The results of statistical comparison are shown as asterisks (*** $p<0.001$; ** $p<0.01$; NS $p>0.05$).

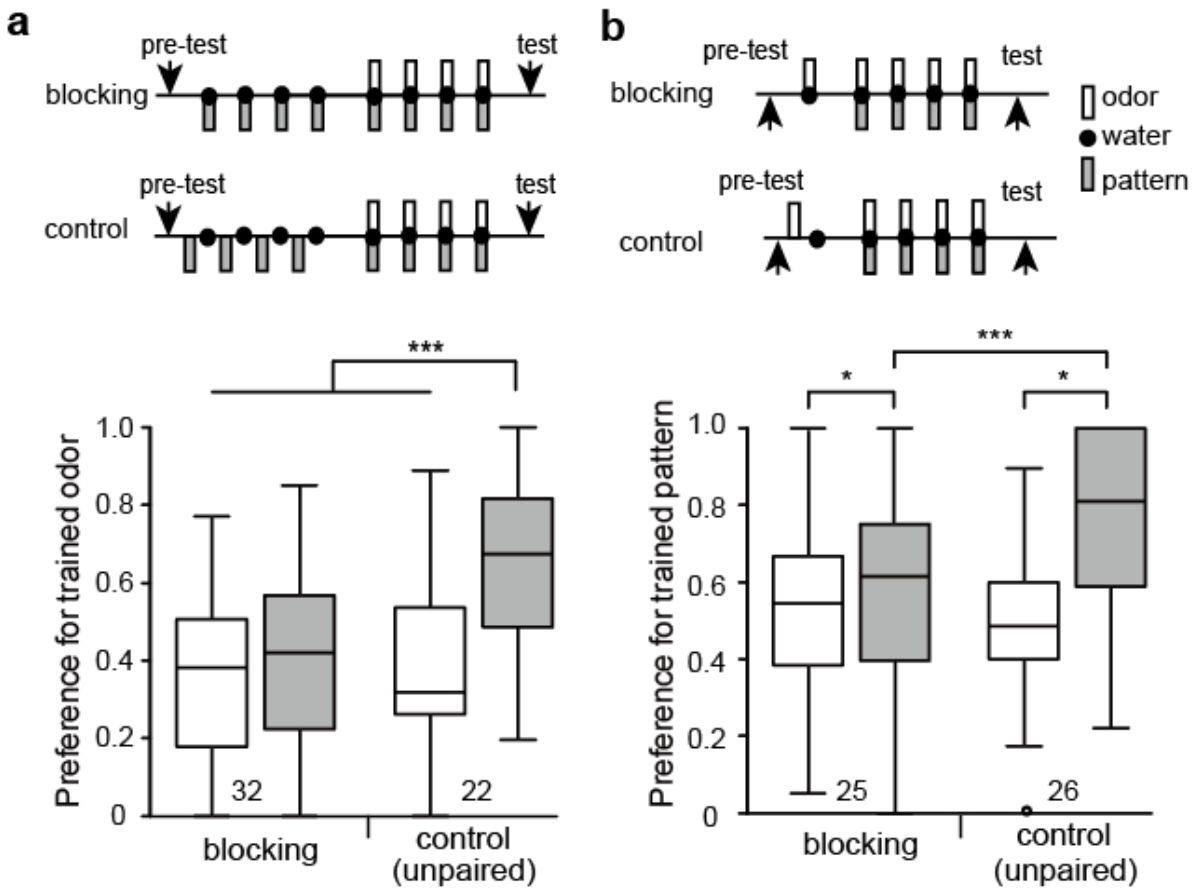


Figure 3. Blocking of visual and olfactory learning. (a) Blocking of olfactory learning. One group of animals (blocking group) was subjected to 4-trial pairing of a pattern with water, and 30 min later the group was subjected to 4-trial pairing of an odor-pattern compound with water US. Another group (unpaired group) was subjected to unpaired presentation of a pattern and water 4 times each, and 30 min later the group was subjected to 4-trial pairing of an odor-pattern compound with water. (b) Blocking of visual learning. One group was subjected to 1-trial pairing of an odor with water and 30 min later subjected to 4-tiral pairing of an odor-pattern compound with water. Another group was subjected to unpaired presentation of an odor and water and 30 min later subjected to 4-trial pairing of an odor-pattern with water. The ITI was 5 min. Relative preference for odors or patterns was tested before and at 30 min after training. PIs for the rewarded odor or pattern before (white bars) and after (gray bars) training are shown as box and whisker diagrams. The number of animals is shown below the boxes. A GLMM was

used to examine relative preferences for the trained odor or pattern before and after training in the blocking and unpaired control groups (Supplemental table S1). The results of statistical comparison are shown as asterisks (*** $p<0.001$; * $p<0.05$; NS $p>0.05$).

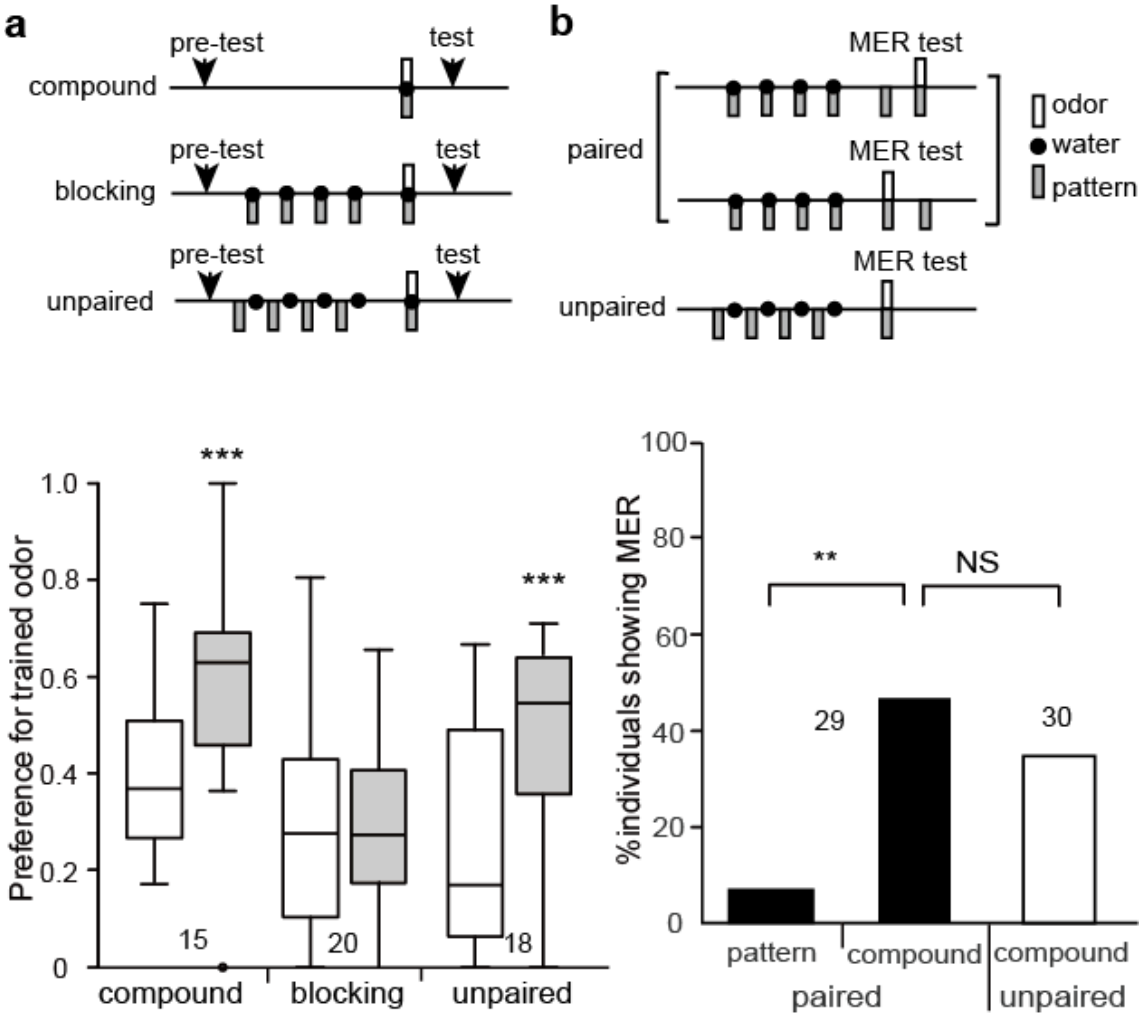


Figure 4. Experiments to discriminate the Rescorla-Wagner model and attentional theory. (a) Blocking with 1-trial compound conditioning. One group of animals (compound group) was subjected to 1-trial pairing of an odor-pattern compound with water. Another group (blocking group) was subjected to 4-trial pairing of a pattern with water, and 30 min later subjected to 1-trial pairing of an odor-pattern compound with water. Another group (unpaired control group) was subjected to unpaired presentation of an odor-pattern compound and water 4 times each, and 30 min later subjected to 1-trial pairing of an odor-pattern compound with water. The ITI was 5 min. Relative odor preferences were tested before and at 30 min after training. PIs for the rewarded odor before and after conditioning are shown as box and whisker diagrams. A GLMM

was used to examine relative preferences for the trained odor before and after training in the compound, blocking and unpaired control groups (Supplemental table S1). (b) Test of attention to an odor presented in odor-pattern compound after pattern conditioning. Two groups of animals were subjected to 4-trial pairing of a pattern with water, and 30 min later one group was tested with MER to the pattern and then to an odor-pattern compound and another group was tested with reversed sequences. Data from the two groups were pooled to measure percentage of MER to the pattern and the compound stimuli because the sequence of tests had no effect. Another group was subjected to unpaired presentation of a pattern and water 4 times each, and 30 min later an odor-pattern compound was presented to test MER. The ITI in phase 1 training was 5 min and the interval between phase 1 and phase 2 training was 30 min. Bars represent percentage of MER to the CS. The number of animals is shown in the figure. McNemar's test was used for comparison of %MER to pattern and compound stimuli. Fisher's exact test was used to compare between groups. For multiple comparisons, Holm's method was used to adjust the significance level. The results of statistical comparison are shown as asterisks (*** $p<0.001$; ** $p<0.01$; NS $p>0.05$).

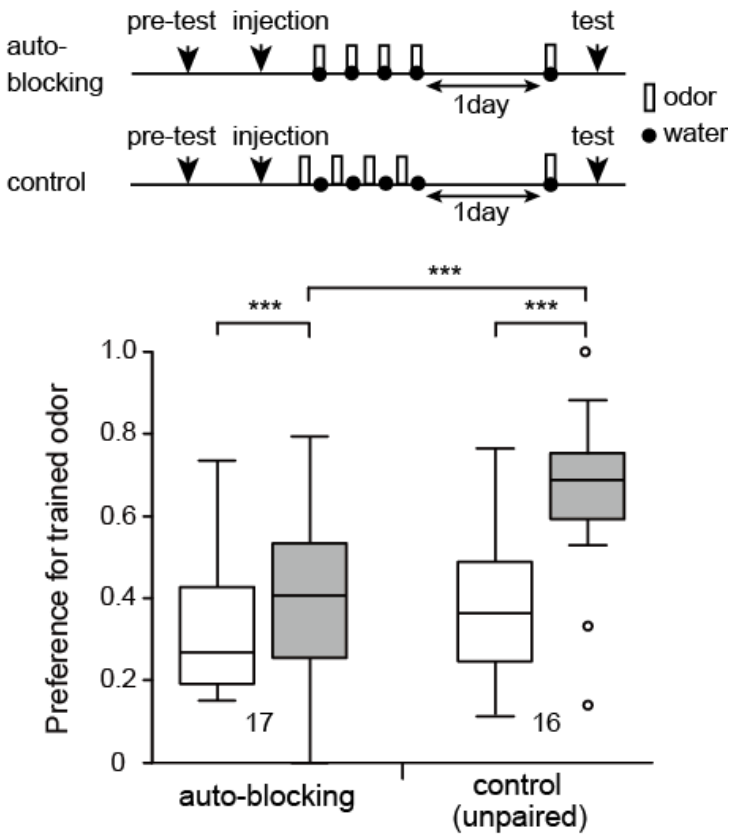


Figure 5. Auto-blocking. Two groups of animals were subjected to injection of 3 μl saline containing 2 μM epinastine. Thirty min later, one group (blocking group) was subjected to 4-trial pairing of an odor with water and another group (control group) was subjected to unpaired presentation of an odor and water 4 times each. The ITI was 5 min. On the next day, both groups were subjected to 1-trial pairing of the odor with water. Relative odor preference was tested before and at 30 min after training. PIs for the rewarded odor before and after training are shown as box and whisker diagrams. The number of animals is shown below the boxes. A GLMM was used to examine relative preferences for the trained odor or pattern before and after training in the blocking and unpaired control groups (Supplemental table S1). The results of statistical comparison are shown as asterisks (*** $p<0.001$).

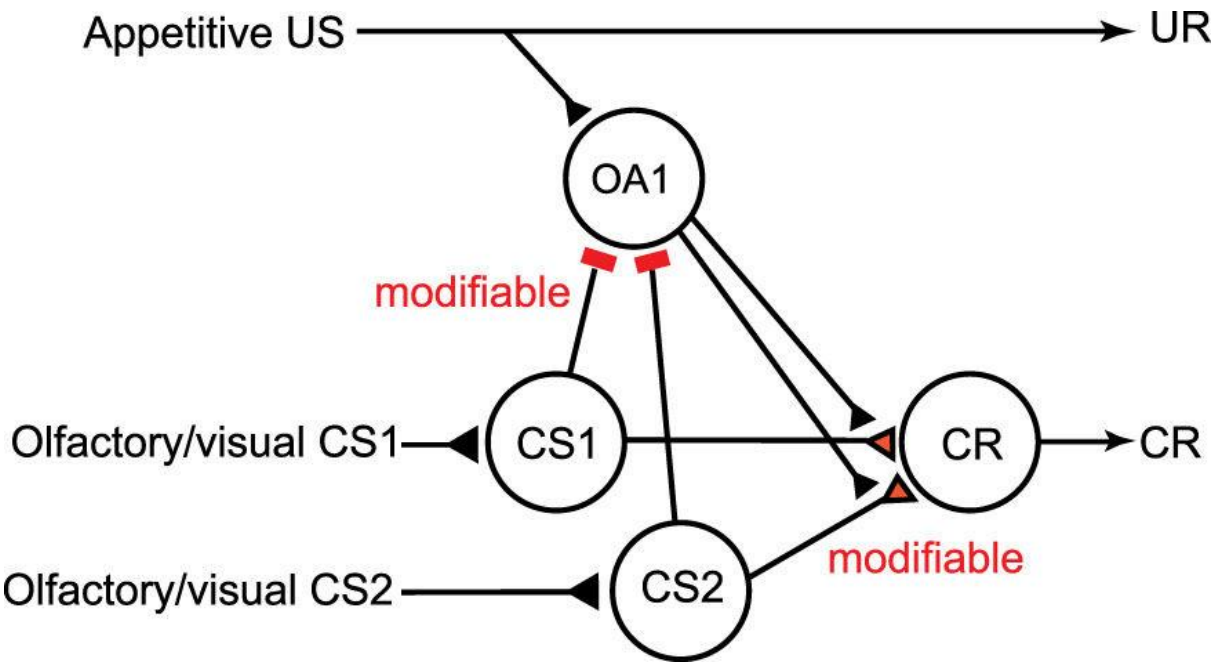


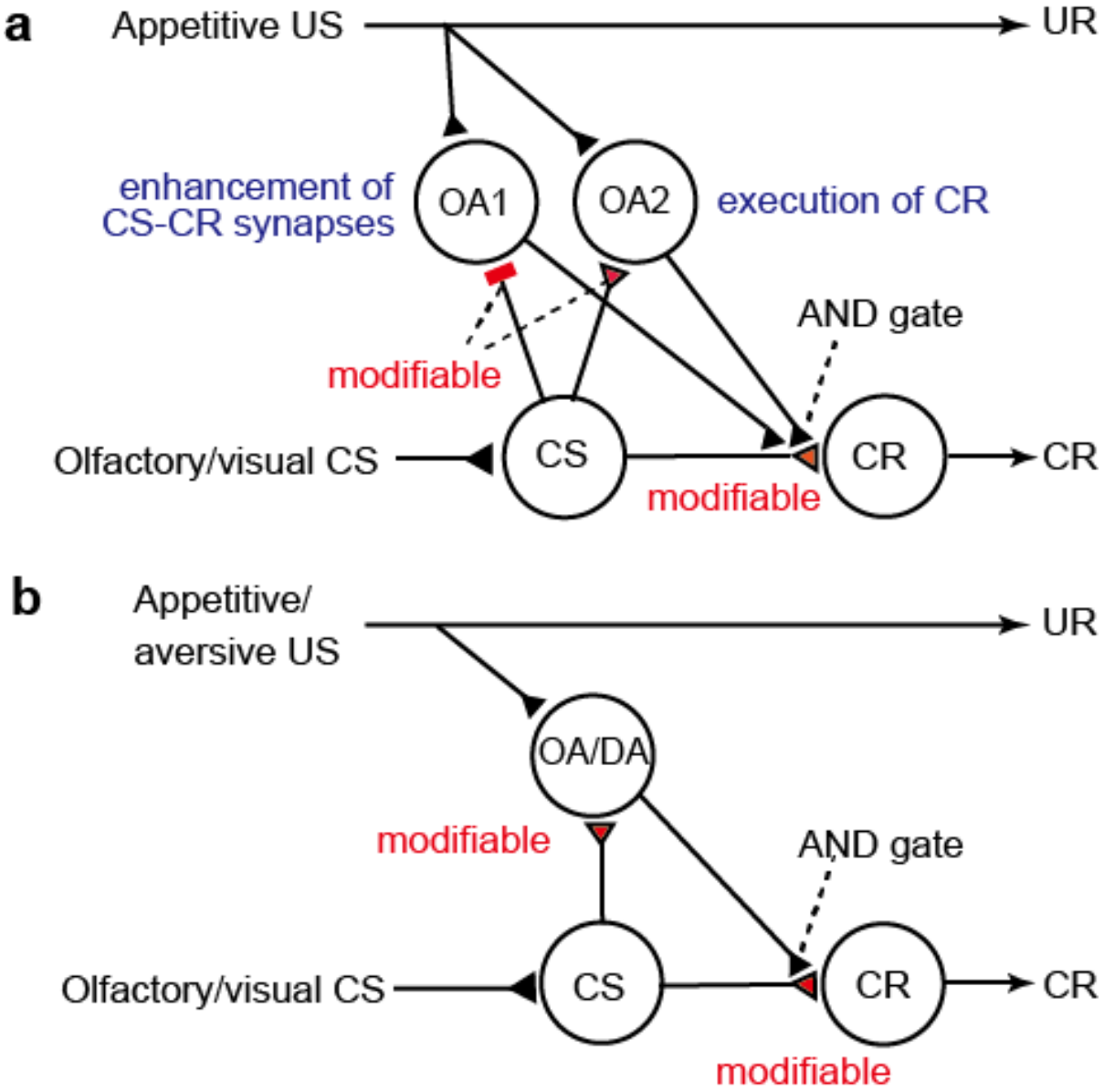
Figure 6. Accounts for blocking by my classical conditioning model. The model shown in Supplementary Fig. S1a assumes that pairing of a stimulus (CS1) with appetitive US leads to an enhancement of inhibitory synapses from "CS1" neurons to "OA1" neurons and that of excitatory synapses from "CS1" neurons to "CR" neurons, in which "CS1" are neurons mediating CS1, "OA1" are a type of octopaminergic (OA) neurons mediating appetitive US and "CR" are neurons whose activation leads to CR. During subsequent pairing of a compound of CS1 and CS2 with US, "OA1" neurons are inhibited by activation of "CS1" neurons and thus responses of "OA1" neurons to US is diminished. As a result, no enhancement of "CS2-OA1" synapses and "CS2-CR" synapses occur, in which "CS2" are neurons mediating CS2. Thus, no learning of CS2 occurs. "OA2" neurons in the model are not shown for simplicity.

Table 1. Procedures and results of blocking experiment.

Group	Phase 1 training	Phase 2 training	Results: Learning of Y?	Figures
Compound	-	XY+	Yes	Figs. 2, 4
Blocking	X+	XY+	No	Figs. 3, 4
Unpaired (control for blocking)	X / +	XY+	Yes	Figs. 3, 4
Auto-blocking	Y+ (under epinastine)	Y+	No	Fig. 6
Unpaired (control for auto-blocking)	Y / + (under epinastine)	Y+	Yes	Fig. 6

XY+: a compound of stimulus X and stimulus Y is paired with appetitive US; (+): reward; Y / +: unpaired presentation of stimulus Y and reward. In most experiments, X is a visual pattern (P) and Y is an odor (O) (Fig. 1a); in other experiment, the stimulus arrangement is reversed (Fig. 1b).

Supplemental Figures and Tables



Supplementary Figure S1. Models of classical conditioning in crickets. (a) A new model of the roles of OA neurons in appetitive conditioning to match the prediction error theory. In the model, I assumed "OA1" neurons that govern enhancement of "CS-CR" synapses (but not execution of CR), in addition to "OA2" neurons that govern execution of CR or memory retrieval (but not enhancement of "CS-CR" synapses). OA2 neurons but not OA1 neurons govern the "AND gate". The "OA1" neurons are assumed to receive silent or very weak inhibitory synapses from "CS" neurons before training, which are strengthened by CS-

US pairing as following Hebb synaptic rule[1]. Recent study in mammals suggests that inhibitory input from GABA neurons is involved in calculation of prediction error in dopamine neurons[2], thus this assumption is not unnatural. During training, "OA1" neurons receive excitatory synaptic input representing actual US and inhibitory input from "CS" neurons representing US predicted by CS, and thus their activities represent US prediction errors, thereby allowing US prediction error signals to govern enhancement of synaptic transmission. (b) A model of appetitive conditioning, which is proposed to account for findings that OA or DA receptor antagonists impair learning and execution of conditioned response (or memory retrieval) in appetitive or aversive conditioning, respectively[3]. The model assumes that (1) "CS" neurons (which may represent intrinsic neurons of the mushroom body) that convey signals about CS make silent or weak synaptic connections with dendrites of "CR" neurons (which may represent efferent (output) neurons of the mushroom body lobe), activation of which leads to a conditioned response (CR), but these synaptic connections are silent or very weak before conditioning, (2) OA or DA neurons ("OA/DA" neurons), which convey signals for appetitive or aversive US, respectively, make synaptic connections with axon terminals of "CS" neurons, (3) "CS" neurons also make silent synaptic connection with "OA/DA" neurons (which might not be monosynaptic), (4) the efficacy of the synaptic transmission from "CS" neurons to "CR" neurons and to "OA/DA" neurons is strengthened by coincident activation of "CS" neurons and "OA/DA" neurons during appetitive or aversive conditioning and (5) after conditioning, activation of "CS" neurons activates "OA" neurons and coincident activation of "CS" neurons and "OA/DA" neurons is needed for activation of "CR" neurons (AND gate) and for production of conditioned responses to CS. UR: unconditioned response.

1. Hebb, DO., The organization of behavior: A neuropsychological theory. New York, Wiley & Sons. (1949).

2. Eshel, N. et al. Arithmetic and local circuitry underlying dopamine prediction errors. Nature

525, 243–246 (2015).

3. Mizunami, M., et al. Roles of octopaminergic and dopaminergic neurons in appetitive and aversive memory recall in an insect. BMC Biol. 7, 46 (2009).

Supplemental table S1. Summary of generalized linear mixed model (GLMM).

a. Compound conditioning experiment for olfactory learning (Fig. 2a)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	-0.7637	0.1403	-5.442	$5.26 * 10^{-8}$
Test	1.111	0.08914	12.46	$2 * 10^{-16}$
Training	0.02090	0.2411	0.087	0.931
Test * Training	-0.07668	0.1630	-0.471	0.638

b. Compound conditioning experiment for visual learning (Fig. 2b)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	-0.2973	0.2369	-1.255	0.210
Test	1.692	0.1715	9.867	$2 *10^{-16}$
Training	0.4404	0.3398	1.296	0.1950
Test * Training	-0.7341	0.2407	-3.050	$2.29 *10^{-3}$

c. Blocking experiment for olfactory learning (Fig. 3a)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	-0.6902	0.1594	-4.331	$1.48 * 10^{-5}$
Test	0.1254	0.09176	1.367	0.172
Training	0.07515	0.2486	0.302	0.762
Test * Training	1.048	0.1442	7.265	$3.73 * 10^{-13}$

d. Blocking experiment for visual learning (Fig. 3b)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	-0.03000	0.1684	-0.178	0.8586

Test	0.2533	0.1215	2.085	0.0371
Training	-0.07646	0.2356	-0.325	0.7455
Test * Training	0.7824	0.1780	4.396	$1.1 * 10^{-5}$

e. One-trial blocking experiment in olfactory learning (Fig. 4a)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	-0.9833	0.2096	-4.692	$2.7 * 10^{-6}$
Test	0.1187	0.1256	0.945	0.344
Training (compound)	0.4629	0.3188	1.452	0.147
Training (unpaired)	-0.09313	0.3067	-0.304	0.761
Test * Training (compound)	1.013	0.1846	5.487	$4.09 * 10^{-8}$
Test * Training (unpaired)	0.9037	0.1786	5.060	$4.19 * 10^{-7}$

f. Auto-blocking experiment with epinastine in olfactory learning (Fig. 5)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	-0.7135	0.1816	-3.930	$8.50 * 10^{-5}$
Test	0.4046	0.1126	3.594	$3.25 * 10^{-4}$
Training	0.08362	0.2606	0.321	0.7483
Test * Training	0.7862	0.1662	4.732	$2.23 * 10^{-6}$

By using a GLMM, effects of the test situation (before or after training), training procedure (compound conditioning or control procedure in a and b, blocking or control procedure in c and d, blocking, compound or control procedure in e, auto-blocking or control procedure in f) and interaction between the test and training on relative preference for the trained odor or pattern were evaluated. The estimate

for the intercept indicates the estimate before training in the compound group (a and b), blocking group (c, d and e) or auto-blocking group (f).

Supplementary Table S2. Information coded in the responses of OA1 and OA2 neurons in my model.

Stimulus	OA1		OA2	
	Before training	After training	Before training	After training
US	1(US)	1 (US)	1(US)	1 (US)
CS	0	0 [-1 (-USP)]*	0	1 (USP)
CS+US	1 (US)	0 (USPE)	1 (US)	1(US+USP)

Responses of OA1 and OA2 neurons in the model shown in Fig. 5B to US (reward), CS, and paired presentation of CS and US before and after conditioning. After completion of training, OA1 neurons that govern enhancement of synaptic transmission underlying conditioning exhibit no responses to paired CS-US presentation, and thus no further enhancement of synaptic transmission occurs. USP: US prediction; USPE: US prediction error. Responses are indicated as all or none (1 or 0). *Negative value in the parentheses indicates inhibitory synaptic input.

Chapter II

Roles of dopamine neurons in mediating the prediction error in aversive learning in insects

Note: This chapter includes materials modified from

Terao, K. & Mizunami, M. Roles of dopamine neurons in mediating the prediction error in aversive learning in insects. Sci. Rep. 7, 14694 (2017).

Abstract

In associative learning in mammals, it is widely accepted that the discrepancy, or error, between actual and predicted reward determines whether learning occurs. The prediction error-based learning theories has been proposed to account for the finding of a blocking phenomenon, in which pairing of a stimulus X with an unconditioned stimulus (US) could block subsequent association of a second stimulus Y to the US when the two stimuli were paired in compound with the same US. Evidence for the prediction error-based learning, however, has been imperfect since blocking can also be accounted for by competitive theories. I recently reported blocking in classical conditioning of an odor with water reward in crickets. I also reported an "auto-blocking" phenomenon in appetitive learning, which supported the prediction error-based learning. The presence of auto-blocking also suggested that octopamine neurons mediate reward prediction error signals. Here I show that blocking and auto-blocking occur in aversive learning to associate an odor with salt water (US) in crickets, and my results suggest that dopamine neurons mediate aversive US prediction error signals. I conclude that the prediction error-based learning mechanisms mediate both appetitive learning and aversive learning in insects.

Introduction

Associative learning allows animals to adapt to various environments by acquiring knowledge on events in their environments. Based on the knowledge, animals find suitable food, avoid toxic food and escape from predators. Thus, both appetitive learning and aversive learning are essential for survival of animals. Many efforts have been made to elucidate learning rules governing associative learning in mammals[1, 2], but whether appetitive learning and aversive learning are ruled by the same general principles remains unclear.

In associative learning in mammals, it is widely accepted that the discrepancy, or error, between the actual unconditioned stimulus (US) and predicted US determines whether learning occurs when a stimulus is paired with the US[1, 2]. This theory stems from the finding of a "blocking" phenomenon by Kamin[3]. He observed in rats that a stimulus X that had been paired previously with a US could block subsequent association of a second stimulus Y to the US when the two stimuli were paired in compound with the same US (XY+ training, see Table 1). Kamin[3] argued that no learning of stimulus Y occurs since the US was fully predicted by stimulus X and argued that surprise is needed for learning. This proposition was formulated into the prediction error-based learning theories including Rescorla-Wagner model[4] and attentional theories[5, 6] (see General introduction). Subsequent electrophysiological studies suggested that dopamine (DA) neurons in the midbrain convey reward prediction error signals[1].

Evidence for the prediction error-based learning theories, however, has been imperfect since blocking can also be accounted for by theories other than the prediction error-based learning theories. One of the most dominant alternative theories is comparator hypothesis(or retrieval theory)[7], which account for blocking by competition between X and Y stimuli. Evidence to convincingly refute alternative theories has been lacking[8-10].

I previously reported blocking in appetitive associative learning in crickets[11]. Moreover, based on the previous studies that octopamine (OA) neurons play critical roles in appetitive learning in crickets[12-19], I demonstrated that when a stimulus X was paired with water

(appetitive US) under the condition of administration of an OA receptor antagonist, in which no learning of X occurs, subsequent learning of X was blocked in training to associate the stimulus X with the US given after recovery from the effect of the antagonist[11]. This "auto-blocking" can be accounted for by the prediction error-based learning theories since if blockade of OA-ergic transmission impairs learning but not formation of the prediction of the US by stimulus X, no learning of stimulus X should occur in subsequent training. This "auto-blocking" phenomenon cannot be accounted for by any of the competitive theories to account for blocking since it occurs without stimulus competition. Therefore, demonstration of blocking and auto-blocking phenomena in the same learning paradigm in the same species provided rigorous evidence for the prediction error-based learning in appetitive conditioning. In addition, the results of an auto-blocking experiment suggested that OA neurons mediate reward prediction error signals in crickets. However, rigorous evidence to show the prediction error-based learning in aversive learning has been still lacking.

In the present study, I investigated whether blocking and auto-blocking occur in aversive learning in crickets. I have shown that DA neurons play critical roles in aversive learning in crickets[12-19], as has been reported for other invertebrates[20-23]. I obtained evidence of blocking in conditioning to associate an odor or pattern with NaCl solution (aversive US) in crickets. I also found that "auto-blocking" occurs in aversive learning, that is, no learning of an odor X occurs in training to associate X with aversive US when the training is preceded by the same training under the condition of administration of a DA receptor antagonist. This blockade of learning was accounted for by the prediction error-based learning theories but not by an alternative theory to account for blocking since no cue competition is involved.

Results

Effects of compound conditioning

Since a blocking experiment requires conditioning of two stimuli presented at the same time, I first investigated whether crickets exhibit such compound conditioning. I used odor-pattern compound conditioning (OP+ conditioning), in which a compound stimulus consisting of an odor (O) and a visual pattern (P) is paired with a 20% NaCl solution (aversive US) (+), and I investigated whether OP+ training leads to learning of the odor or the visual pattern (Fig. 1 and Table 1). One group of animals (compound group) was subjected to 2-trial OP+ training and another group (control group) was subjected to 2-trial olfactory conditioning (O+ conditioning). Relative preference for the odor used in training compared to the control odor was tested before and at 20 min after training in both groups. The results are shown in Fig. 2a. I used a generalized linear mixed model (GLMM) to evaluate the data (see Methods). Both the compound group and control group exhibited significantly decreased preference for the conditioned odor after training compared to that before training (test term, $p = 8.68*10^{-10}$, $z = -6.132$, see Supplemental table S1). The preference for the conditioned odor after training in the compound group did not significantly differ from that in the control group (test * training term, $p = 0.141$, $z = 1.471$). The results showed that learning was achieved in the compound group as in the control group. I thus conclude that odor-pattern compound conditioning leads to conditioning of the odor.

Next, I investigated whether OP+ training leads to learning of the visual pattern. One group of animals (compound group) was subjected to 8-trial OP+ training and another group (control group) was subjected to 8-trial visual conditioning (P+ conditioning). Relative preference for the pattern used in training compared to the control pattern was tested before and at 20 min after training in both groups. The results are shown in Fig. 2b. Both the compound group and control group exhibited significantly decreased preference for the conditioned pattern after training compared to that before training (test term, $p = 7.97 *10^{-15}$, $z = -7.768$). The results showed that learning was achieved in both the compound group and the control group. In addition, I

observed that the preference for the conditioned pattern after training in the compound group was significantly less than that in the control group (test * training term, $p = 1.35 * 10^{-4}$, $z = 3.818$). These unexpected results are discussed in a later section.

Demonstration of blocking

I next studied whether blocking occurs in aversive learning in crickets. At first, I investigated whether blocking of olfactory learning occurs. One group of crickets (blocking group) was subjected to 8-trial P+ training and then 2-trial OP+ training (Table 1). Another group (control group) was subjected to unpaired presentations of a visual pattern and aversive US (P / + training) 8 times each and then 2-trial OP+ training. The results are shown in Fig. 3a. The preference for the trained odor after training in the control group was significantly less than that before training in the same group and was also significantly less than that before or after training in the blocking group (test * training term, $p = 0.00148$, $z = -3.179$). The results showed that conditioning was achieved in the control group but not in the blocking group.

I next studied whether blocking of visual pattern learning occurs. One group of crickets (blocking group) was subjected to 2-trial O+ training and then 8-trial OP+ training. Another group (control group) was subjected to unpaired presentations of an odor and aversive US (O / + training) 2 times each and then 8-trial OP+ training. The results are shown in Fig. 3b. The preference for the trained pattern after training in the control group was significantly less than that before training in the same group and was also significantly less than that before or after training in the blocking group (test * training term, $p = 3.6 * 10^{-8}$, $z = -5.509$). The results showed that conditioning was achieved in the control group but not in the blocking group. The results indicate that blocking occurs in visual learning.

A neural circuit model of classical conditioning that matches the prediction error-based learning theories I previously proposed a neural circuit model for appetitive learning that

matches the prediction error-based learning theories[11]. The model was designed to represent neural circuits in lobes of the MB, which is known to play critical roles in learning[20,21], and was based on findings that OA neurons play critical roles in appetitive learning in crickets[12-17,19]. Here I propose a model of aversive learning that matches the prediction error-based learning theories (Fig. 4), in which I focused on the roles of DA neurons in aversive learning[12-19]. For complete description of my model, see Supplementary figure S2.

In the model shown in Fig. 4a, "DA" neurons (assuming DA neurons projecting to the lobes of the MB) are assumed to receive inhibitory synapses from "CS" neurons (assuming Kenyon cells of the MB), the efficacy of which is strengthened by conditioning. In pairing of an olfactory CS with a sodium chloride US, "DA" neurons receive excitatory input representing actual US and inhibitory input representing predicted US by the CS, and their responses thus represent US prediction error signals. Hence, US prediction error signals govern enhancement of synaptic transmission that underlies conditioning. How the model accounts for blocking is shown in Fig.4b (for an explanation, see legends). To better account for the model, information coded by "DA" neurons before and after training is shown in Supplemental table S2.

Demonstration of auto-blocking

My model predicts that blockade of synaptic transmission from DA neurons by a DA receptor antagonist (flupentixol[24]) during Y+ training impairs learning of Y but not formation of aversive US prediction by Y since, assuming that the antagonist impairs enhancement of "CS-CR" synapses but not that of "CS-DA" synapses (see Fig. 4a), subsequent Y+ training given after recovery from the effect of the antagonist should produce no learning. This effect is termed "auto-blocking", because learning of Y is blocked by US prediction by Y itself, not by X in the case of blocking. I previously reported such an auto-blocking phenomenon in appetitive learning in crickets by using OA receptor antagonist (epinastine)[11].

I tested whether auto-blocking occurs in aversive learning in crickets. One group of animals

(auto-blocking group) was injected with flupentixol into the head hemolymph and 30 min later the group was subjected to 6-trial O+ training. The dose of flupentixol was determined based of previous studies[15-17]. On the next day, the group was subjected to 2-trial O+ training. Another group (control group) was subjected to unpaired presentation of the odor and aversive US (O / + training) 6 times each under the condition of application of flupentixol and then was subjected to 2-trial O+ training the next day. The results are shown in Fig. 5. The preference for the trained odor after training in the control group was significantly less than that before training in the same group and was significantly less than that before or after training in the auto-blocking group (test * training term, $p = 0.00144$, $z = -3.186$). The results showed that learning was achieved in the control group but not in the auto-blocking group and indicate that auto-blocking occurs in aversive learning in crickets.

Previous studies showed that octopamine receptor antagonist (epinastine) does not impair aversive learning[13-17], and here I performed an experiment to confirm that epinastine does not lead to auto-blocking in aversive learning. One group of animals was injected with epinastine into the head hemolymph and 30 min later the group was subjected to 6-trial O+ training. On the next day, the group was subjected to 2-trial O+ training. The results are shown in Supplemental Figure S1. The preference for the trained odor after training was significantly less than that before training (test term, $p = 7.23 * 10^{-4}$, $z = -3.381$), indicating that learning was successful. I conclude that DA receptor antagonist but not OA receptor antagonist leads to auto-blocking of aversive learning.

Discussion

I obtained convincing evidence for the prediction error-based learning in aversive learning. I demonstrated, at first, that a blocking phenomenon occurs in aversive learning in crickets, i.e., no learning of Y occurred by XY+ training when the training was preceded by X+ training with X and Y being either visual or olfactory stimulus. Then I proposed a neural circuitry model of aversive learning, in which a previous model of aversive learning[15] was modified to match the prediction error-based learning theories. My aversive learning model (Fig. 4) was a counterpart of the appetitive learning model I proposed previously[11] and predicted an "auto-blocking" phenomenon, in which no learning of X occurs by X+ training when the training is preceded by X+ training under the condition of administration of a DA receptor antagonist, and I indeed observed this phenomenon in olfactory learning. The results of the auto-blocking experiment showed the validity of the prediction error-based learning: To my knowledge, alternative theories to account for blocking other than the prediction error-based learning theories, including naive attentional theories (see chapter 1) and retrieval theories[7] (or comparator hypothesis), assume cue competition between X and Y to account for blocking, thus fail to account for auto-blocking. Demonstration of blocking and auto-blocking phenomena in aversive learning (this study) and in appetitive learning[11] in the same species provides rigorous evidence for the prediction error-based learning in both appetitive and aversive forms of olfactory learning in crickets. Demonstration of auto-blocking of visual learning remains for my future subject.

Previous reports on blocking in aversive learning in animals

Blocking has been reported in various systems of aversive learning in vertebrates and invertebrates. A blocking phenomenon was first demonstrated in classical conditioning of tone and light compound stimuli with electric shock US in rats[3]. Evaluation of this learning paradigm led to proposals of the Rescorla-Wagner model[4], attentional theory[5, 6] and retrieval theory[7].

Blocking in aversive learning has also been reported in mollusks, in which odor, light or tactile stimulus was paired with bitter taste, electric shock or other aversive US[25-27]. Some researchers attempted to discriminate learning theories to account for blocking in aversive conditioning, but convincing evidence to discriminate among all different theories has not been reported[8-10].

My blocking and auto-blocking experiments produced evidence for the prediction-error based learning theories, and discrimination among them is an important future subject. As discussed in chapter 1, one-trial blocking is suitable experiment for the subject.

I observed that the effect of compound conditioning of a visual pattern and an odor was significantly more than that of conditioning of a visual pattern alone in both appetitive and aversive learning (Fig. 2b in chapter 1 and 2), indicating that simultaneous presentation of an olfactory cue facilitated conditioning of a visual cue. This is an unexpected observation since I did not find such an effect in olfactory conditioning[11]. Similar results and possible accounts are proposed in mammals[38, 39]. Comparative studies between mammals and crickets should be needed for this subject.

Roles of dopamine neurons in mediating prediction error signals about aversive US

DA neurons are thought to convey reinforcement signals in many systems of associative learning in insects and mammals. In the fruit-fly *Drosophila*, it has been suggested that different classes of DA neurons projecting to the lobes of the MB mediate reinforcement signals in aversive learning and appetitive learning[20, 21]. In honey bees, as in crickets, it has been suggested that DA neurons convey reinforcement signals in aversive learning[22], whereas OA neurons convey reinforcement signals in appetitive learning[28, 29]. However, the exact nature of signals that DA or OA neurons convey in learning has not been characterized in any insects. Future electrophysiological studies on activities of DA neurons during conditioning are needed to clarify this issue.

In mammals, there is evidence that midbrain DA neurons mediate prediction error signals

in appetitive learning[1, 2, 30, 31], but the roles of DA neurons in aversive learning remain controversial. Some researchers have suggested that midbrain DA neurons participate in aversive learning[32, 33] and convey aversive US prediction error[34], but other researchers have argued that midbrain neurons mediating aversive signals may not be DAergic[31, 35, 36]. To what extent the roles of DA neurons in associative learning are conserved between insects and mammals remains for a fascinating research subject.

Are there interactions between neurons mediating prediction errors about reward and aversive US?

I suggest that OA and DA neurons convey prediction error signals in appetitive learning and aversive learning, respectively, in crickets and an important future subject is to investigate whether OA and DA neurons independently process reward and aversive US prediction error signals, respectively, or whether these neurons tightly interact to integrate reward and aversive US prediction error signals and to form a unified system to mediate value prediction error signals in insects. Previous studies demonstrated that intervention of DA-ergic transmission by DA receptor antagonists or by knockdown or knockout of genes that code for a type of DA receptor by RNAi or by the CRISPR/cas9 system impaired aversive learning but did not affect appetitive learning, whereas intervention of OA-ergic transmission impaired appetitive learning but not aversive learning[12-19]. In this study, I showed that DA receptor antagonist but not OA receptor antagonist leads to auto-blocking of aversive learning. The results indicate that the OA reward system and DA aversive US system can act independently when appetitive learning and aversive learning occur independently. Those studies, however, do not exclude the possibility that DA and OA neurons interact in a situation in which a stimulus is associated with appetitive and aversive stimuli. A similar issue has been discussed in mammals. Some researchers have suggested that separate classes of midbrain neurons mediate prediction error signals about reward and aversive US[31, 35, 36], whereas other researchers have proposed that a single class of

DA neurons integrates reward and aversive US signals to encode value prediction error signals[34]. Further investigations in insects may help to better clarify this issue.

I conclude that insects predict future biologically significant events by appetitive and aversive associative learning and that DA neurons mediate prediction error signals in aversive learning. Neural circuitry mechanisms for computation of the prediction error is a fascinating subject, and insects should emerge as pertinent models in which to elucidate this important subject.

Methods

Insects. Adult male crickets, *Gryllus bimaculatus*, at 1 week after the imaginal molt were used. Before the experiment, animals were placed individually in beakers and deprived of drinking water for 4 days to enhance their motivation to search for water.

Olfactory and Visual Conditioning Procedures. I used classical conditioning and operant testing procedures described previously[11, 37] (Fig. 1). In olfactory conditioning, maple or vanilla odor (conditioned stimulus, CS) was paired with NaCl solution (aversive US). In visual conditioning, a white-center and black-surround pattern (white-center pattern) was paired with 20% NaCl solution. The outer diameter of the pattern was 4 cm and that of the while center pattern was 3 cm. In the compound conditioning, an odor and a white-center pattern were presented simultaneously (compound CSs) and were paired with NaCl solution. A syringe was used to present the CS and US to each cricket. The syringe contained NaCl solution as US, and at its needle, a filter paper soaked with a drop of odor essence was attached as olfactory CS, and/or a white-center pattern was attached as visual CS (Fig. 1). For a conditioning trial, an odor was approached to the antennae (within 1-2cm) or a visual pattern was approached to the head of the cricket (within 2-3 cm) and held for 3 sec, and then a drop of NaCl solution was attached to its mouth. For an unpaired trial, an odor or a visual pattern was approached to the antennae or the head and held for 3 sec, and 2.5 min later, a drop of NaCl solution was attached to its mouth by another syringe. In all pairing experiments, the intervals between the trials (inter-trial intervals, ITIs) were 5 min. After olfactory or compound conditioning trials, the air in the beaker was ventilated.

Preference Tests. Odor and pattern preference tests were carried out as described previously[11, 37]. All groups were tested with relative preference between the maple odor and vanilla odor from 1 day to 1 hour before conditioning and 20 min or 1 day after conditioning. The test

apparatus consisted of waiting chambers and a test chamber. In the chamber to test odor preference, the floor had two holes that connected the chamber with two cylindrical containers that contained a filter paper soaked with either a drop of maple (NARIZUKA corporation, Tokyo) or vanilla essence (Kyouritsu-foods, Tokyo) diluted with 5 drops of water and was covered with a fine gauze net (Fig. 1). Three containers were mounted on a rotative holder, and two of the three containers could be located simultaneously beneath the holes of the test chamber. In the apparatus for the pattern preference test, two white-center patterns and one black-center pattern were displayed on a gray sliding wall at the end of the test chamber, and a white-center pattern and a black-center pattern could be presented at the same time (Fig. 1). Before testing, a cricket was transferred to the waiting chamber and left for about 4 min to become accustomed to the surroundings. Then the cricket was allowed to enter the test chamber and the test started. Two min after the test had started, the relative positions of the odor sources or patterns were changed by rotating the container holder or sliding the wall. The preference test lasted for 4 min. I considered that the cricket visited an odor source or a pattern when the cricket probed the top net or the pattern with its mouth or palpi. The time that the cricket visited each odor sources or patterns was recorded cumulatively for each seconds. If the total visiting time of a cricket to odor sources was less than 10 sec, I considered that the animal was less motivated, possibly due to a poor physical condition, and the data were rejected. In the present experiments, about 15% animals were rejected in each test.

Pharmacology. Crickets were injected with 3 μl of saline containing 100 μM flupentixol or 2 μM epinastine (Sigma-Aldrich, Tokyo) into the head hemolymph 30 min before the training. The estimated final concentration after circulation is 350 nM for flupentixol and 7.0 nM for epinastine[11, 12].

Statistical Analysis. Relative preference for the conditioned odor compared with the control

odor was determined as the proportion of time spent visiting the conditioned odor in the total time spent visiting the two odors. I measured the search time with the accuracy of seconds. In previous studies, non-parametric statistical tests are used for evaluation of the relative preference. I used a GLMM with a binomial distribution of the relative preference, determined by the search time data sampled for each second, and logit link function. I included the test condition (test before or after training), training procedure and the interaction term (test * training) as fixed effects in the GLMM, with the training and test terms being categorical variables. Individual cricket was used as a random effect, allowing the random intercept. I used R (ver. 3. 3. 1) and lme4 (ver. 1.1.12) packages for statistical analysis. I refer to as "significantly different" if p-values in the Wald statistical test in the GLMM were $p<0.05$.

References

1. Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87-115 (2006).
2. Schultz, W. Updating dopamine reward signals. *Curr. Opin. Neurobiol.* **23**, 229-238 (2013).
3. Kamin, L. [Predictability, surprise, attention and conditioning] *Punishment and aversive behavior* [Campbell, B.A. & Church, R.M. (eds.)] [279-298] (Appleton-Century-Crofts, New York, 1969).
4. Rescorla, R. A. & Wagner, A. R. [A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement] *Classical Conditioning II* [Black, A. & Prokasy, W.R. (eds.)] [64-99] (Academic Press, New York, 1972).
5. Mackintosh, N. J. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol. Rev.* **82,** 276-298 (1975).
6. Pearce, J. M. & Hall, G. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol. Rev.* **87**, 532-552 (1980).
7. Miller, R. R. & Matzel L. D. The comparator hypothesis: a response rule for the expression of associations. *Psychol. Learn. Motiv.* **22**, 51-92 (1988).
8. Pearce, J. M. [Associative learning] *Animal Learning & Cognition* [35-91] (Psychology press, New York, 2008).
9. Miller, R. R., Barnet, R. C. & Grahame, N. J. Assessment of the Rescorla-Wagner model. *Psychol. Bull.* **117**, 363-386 (1995).
10. Mazur, J. E. [Chapter 4: Theories and research on classical conditioning] *Learning and behavior* [75-100] (Pearson education, Boston, 2013).
11. Terao, K., Matsumoto, Y. & Mizunami, M. Critical evidence for the prediction error theory in associative learning. *Sci. Rep.* **5**, 8929 (2015).
12. Unoki, S., Matsumoto, Y. & Mizunami, M. Participation of octopaminergic reward system and dopaminergic punishment system in insect olfactory learning revealed by

pharmacological study. *Eur. J. Neurosci.* **22**, 1409-1416 (2005).

13. Unoki, S., Matsumoto, Y. & Mizunami, M. Roles of octopaminergic and dopaminergic neurons in mediating reward and punishment signals in insect visual learning. *Eur. J. Neurosci.* **24**, 2031-2038 (2006).
14. Nakatani, Y., *et al.* Why the carrot is more effective than the stick: Different dynamics of punishment memory and reward memory and its possible biological basis. *Neurobiol. Learn. Mem.* **92**, 370-380 (2009).
15. Mizunami, M., *et al.* Roles of octopaminergic and dopaminergic neurons in appetitive and aversive memory recall in an insect. *BMC Biol.* **7**, 46 (2009).
16. Matsumoto, Y., Hirashima, D. & Mizunami, M. Analysis and modeling of neural processes underlying sensory preconditioning. *Neurobiol. Learn. Mem.* **101**, 103-113 (2013).
17. Matsumoto, Y., Matsumoto, C. S., Wakuda, R., Ichihara, S. & Mizunami M. Roles of octopamine and dopamine in appetitive and aversive memory acquisition studied in olfactory conditioning of maxillary palpi extension response in crickets. *Front. Behav. Neurosci.* **9**, 230 (2015).
18. Awata, H. *et al.* Knockout crickets for the study of learning and memory: Dopamine receptor Dop1 mediates aversive but not appetitive reinforcement in crickets. *Sci. Rep*. **5**, 15885 (2015).
19. Awata, H. *et al.* Roles of OA1 octopamine receptor and Dop1 dopamine receptor in mediating appetitive and aversive reinforcement revealed by RNAi studies. *Sci. Rep.* **6**, 29696 (2016).
20. Kim, Y. C., Lee, H. G. & Han, K. A. D1 dopamine receptor dDA1 is required in the mushroom body neurons for aversive and appetitive learning in *Drosophila*. *J. Neurosci.* **27**, 7640-7647 (2007).
21. Aso, Y. *et al.* Three dopamine pathways induce aversive odor memories with different stability. *PLoS Genet*. **8**, (2012).

22. Vergoz, V., Roussel, E., Sandoz, J. C. & Giurfa, M. Aversive learning in honeybees revealed by the olfactory conditioning of the sting extension reflex. *PLoS One* **2**, e288 (2007).
23. Klappenbach, M., Maldonado, H., Locatelli, F. & Kaczer, L. Opposite actions of dopamine on aversive and appetitive memories in the crab. *Learn. Mem.* **19**, 73-83 (2012).
24. Mustard, J. A. *et al.* Analysis of two D1-like dopamine receptors from the honey bee Apis mellifera reveals agonist-independent activity. *Mol. Brain Res.* **113**, 67-77 (2003).
25. Sahley, C., Rudy, J. W. & Gelperin, A. An analysis of associative learning in a terrestrial mollusc. *J. Comp. Physiol. A.* **144**, 1-8 (1981).
26. Rogers, R. F. & Matzel, L. D. Higher-order associative processing in *Hermissenda* suggests multiple sites of neuronal modulation. *Learn. Mem.* **2**, 279-298 (1996).
27. Prados, J. *et al*. Blocking in rats, humans and snails using a within-subjects design. *Behav. Process.* **100**, 23-31 (2013).
28. Hammer, M. & Menzel, R. Multiple sites of associative odor learning as revealed by local brain microinjections of octopamine in honeybees. *Learn. Mem.* **5**, 146-56 (1998).
29. Farooqui, T., Robinson, K., Vaessin, H. & Smith, B. H. Modulation of early olfactory processing by an octopaminergic reinforcement pathway in the honeybee. *J. Neurosci*. **23**, 5370-5380 (2003).
30. Steinberg, E. E. *et al*. A causal link between prediction errors, dopamine neurons and learning. *Nat. Neurosci.* **16**, 966-973 (2013).
31. Schultz, W. Neuronal reward and decision signals: From theories to data. *Physiol. Rev*. **95**, 853-951 (2015).
32. Li, S. S. Y. & McNally, G. P. The conditions that promote fear learning: Prediction error and Pavlovian fear conditioning. *Neurobiol. Learn. Mem.* **108**, 14-21(2014).
33. Wenzel, J. M., Rauscher, N. A., Cheer, J. F. & Oleson, E. B. A role for phasic dopamine release within the nucleus accumbens in encoding aversion: A review of the neurochemical literature. *ACS Chem. Neurosci.* **6**, 16-26 (2015).

34. Matsumoto, H., Tian, J., Uchida, N. & Watabe-Uchida, M. Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *Elife* **5**, 1-24 (2016).
35. Fiorillo, C. D. Two Dimensions of Value: Dopamine neurons represent reward but not aversiveness. *Science* **341**, 546-549 (2013).
36. Stauffer, W. R., Lak, A., Kobayashi, S. & Schultz, W. Components and characteristics of the dopamine reward utility signal. *J. Comp. Neurol.* **524**, 1699-1711 (2016).
37. Matsumoto, Y. & Mizunami, M. Temporal determinants of long-term retention of olfactory memory in the cricket *Gryllus bimaculatus*. *J. Exp. Biol.* **205**, 1429-1437 (2002).
38. Durlach, P. J. & Rescorla, R. A. Potentiation rather than overshadowing in flavor-aversion learning: an analysis in terms of within-compound associations. J. Exp. Psychol. **6**, 175–187 (1980).
39. Austen, J. M., Kosaki, Y. & McGregor, A. Within-compound associations explain potentiation and failure to overshadow learning based on geometry by discrete landmarks. J. Exp. Psychol. **39**, 259–272 (2013).

Figures and Tables

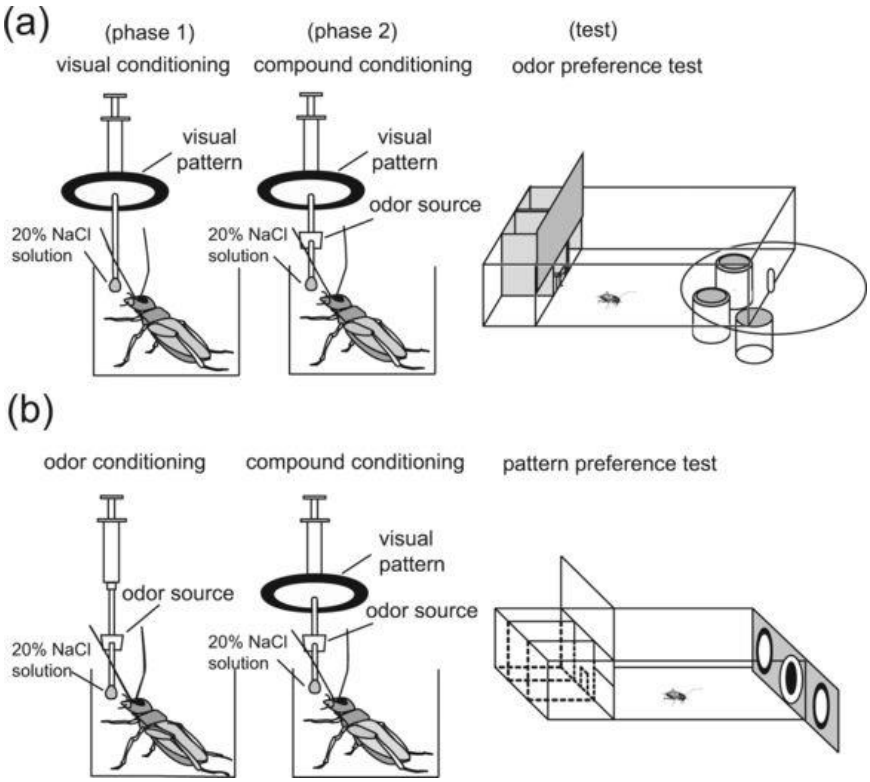


Figure 1. Experimental procedures for blocking of olfactory learning (a) or visual learning (b). (a) A group of water-deprived crickets, individually placed in a beaker, was subjected to pairings of a visual pattern with NaCl solution (P+ training) and then pairings of an odor-pattern compound with NaCl solution (OP+ training). Relative preference for the conditioned odor compared with a control odor was tested before and after training in a test chamber. (b) Another group of crickets was subjected to pairings of an odor with NaCl solution (O+ training) and then pairings of an odor-pattern compound with NaCl solution (OP+ training). Relative preference for the conditioned pattern compared with a control pattern was tested before and after training in a test chamber. The figures were modified from my previous paper[11], licensed by creative commons (CC-BY 4.0: https://creativecommons.org/licenses/by/4.0/legalcode).

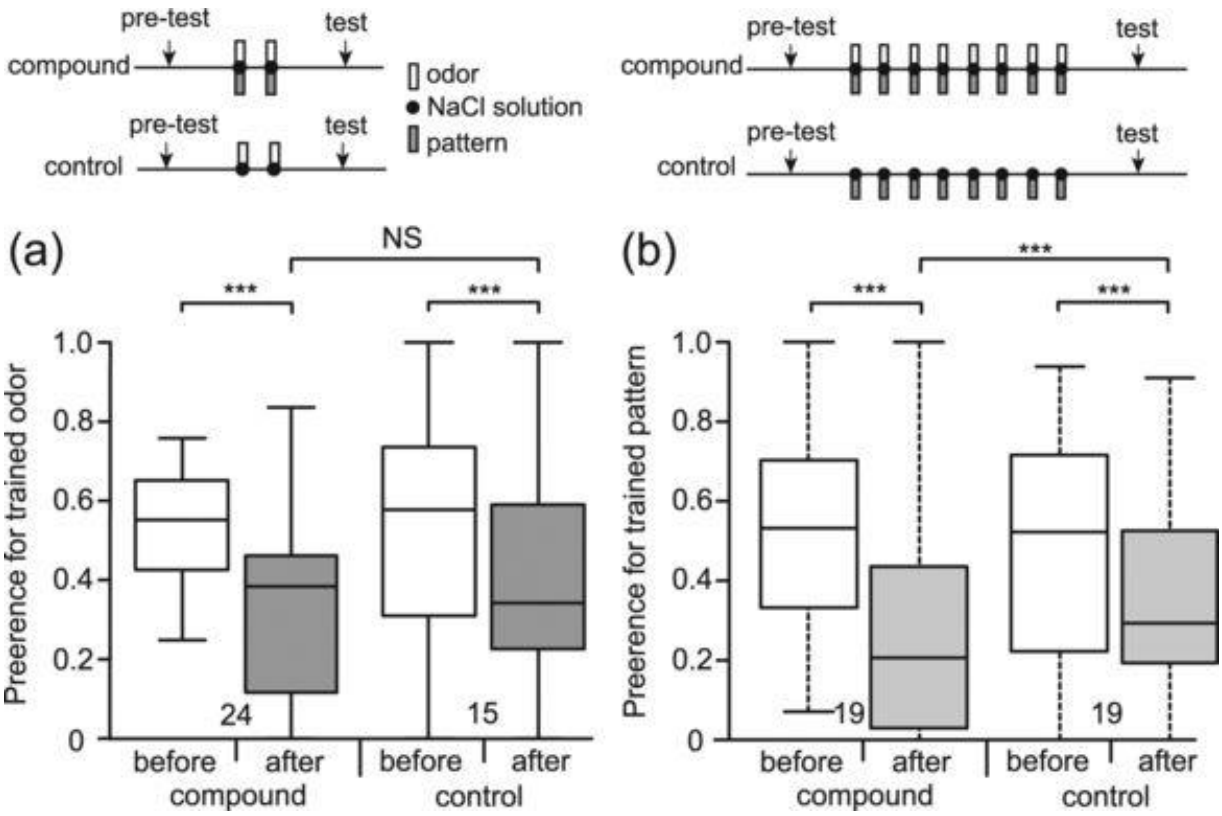


Figure 2. Effect of odor-pattern compound conditioning. (a) One group of animals (compound group) was subjected to 2-trial pairing of an odor-pattern compound with NaCl solution (aversive US) and another group (control group) was subjected to 2-trial pairing of an odor alone with NaCl solution. (b) One group of animals (compound group) was subjected to 8-trial pairing of an odor-pattern compound with NaCl solution and another group (control group) was subjected to 8-trial pairing of a pattern alone with NaCl solution. The inter-trial interval (ITI) was 5 min. Relative preference for the odor or pattern was tested before and at 20 min after training. The experimental procedures are illustrated at the top, and relative preferences for the trained odor or pattern before training (white boxes) and after training (grey boxes) are shown as box and whisker diagrams. The horizontal line in the box is the median, and the box represents the 25-75 percentiles in this and in all following figures. Whiskers extend to extreme values as long as they are within a range of 1.5× box length. The outliers are shown as open circles. The number of animals is shown below the boxes. A GLMM was used to examine relative preferences for the trained odor or pattern before and after training in the

compound and control groups (Supplemental table S1). Statistical significance is shown as asterisks (*** $p<0.001$; NS $p>0.05$).

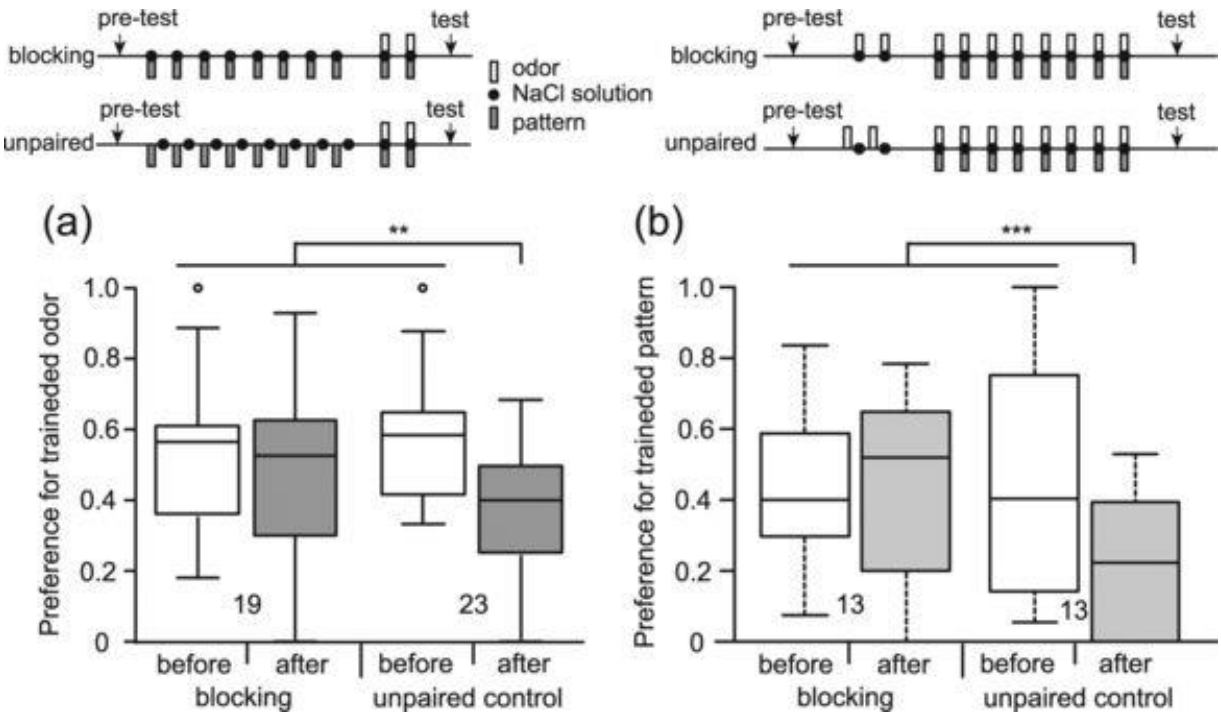


Figure 3. Blocking of olfactory learning (a) or visual learning (b). (a) One group of animals (blocking group) was subjected to 8-trial pairing of a pattern with NaCl solution, and 20 min later the group was subjected to 2-trial pairing of an odor-pattern compound with NaCl solution. Another group (unpaired control group) was subjected to unpaired presentation of a pattern and NaCl solution 8 times each, and 20 min later the group was subjected to 2-trial pairing of an odor-pattern compound with NaCl solution. (b) One group of animals (blocking group) was subjected to 2-trial pairing of an odor with NaCl solution, and 20 min later the group was subjected to 8-trial pairing of an odor-pattern compound with NaCl solution. Another group (unpaired control group) was subjected to unpaired presentation of an odor and NaCl solution 2 times each, and 20 min later the group was subjected to 8-trial pairing of an odor-pattern compound with NaCl solution. The ITI was 5 min. Relative odor or pattern preference was tested before and at 20 min after training. Relative preferences for the trained odor or pattern before (white bars) and after (gray boxes) training are shown as box and whisker diagrams. The number of animals is shown below the boxes. A GLMM was used to examine relative preferences for the conditioned odor or pattern before and after conditioning in the blocking

and unpaired control groups (Supplemental table S1). Statistical significance is shown as asterisks (** $p<0.01$; *** $p<0.001$).

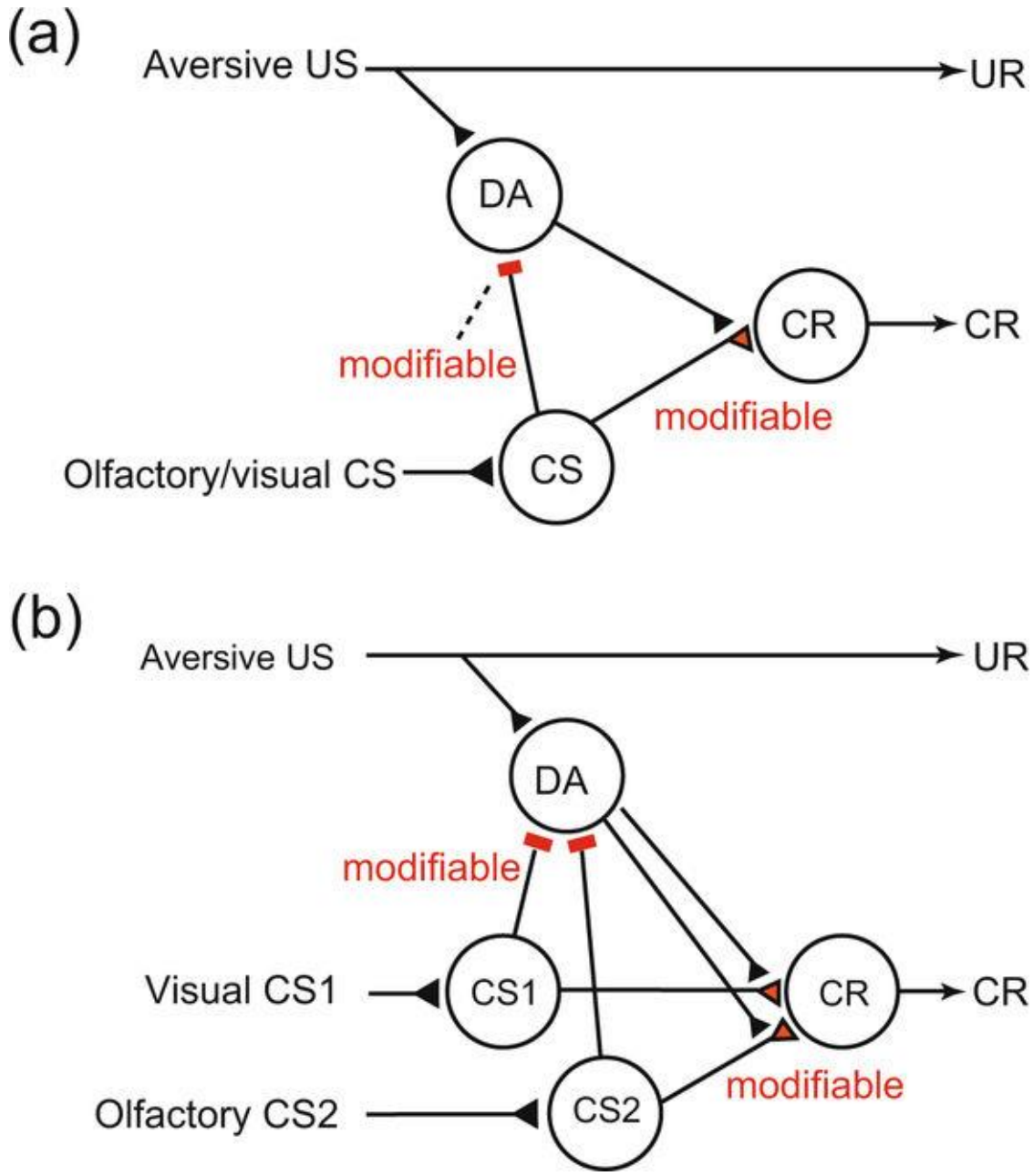


Figure 4. A model of aversive conditioning in crickets. (a) A model for the roles of DA neurons in aversive conditioning to match the prediction error-based learning theories. In the model, "DA" neurons (assuming DA neurons projecting to the mushroom body (MB) lobe) govern enhancement of synaptic transmission that underlies learning. The "DA" neurons are assumed to receive silent or very weak inhibitory synapses from "CS" neurons (assuming Kenyon cells of the MB) before training, which are strengthened by CS-US pairing. During training, "DA" neurons receive excitatory synaptic input representing actual US and inhibitory

synaptic input from “CS” neurons representing US predicted by CS, and thus their activities represent US prediction errors (see Supplemental table S2), thereby allowing US prediction error signals to govern enhancement of synaptic transmission from “CS” neurons to “CR neurons (the latter assuming output neurons from the MB lobe). A complete version of my model is described in Supplemental Figure S2. (b) Accounting for blocking by the model. After a sufficient number of CS1-US pairings, “DA” neurons are inhibited by activation of “CS1” neurons during pairing of a compound of CS1 and CS2 with US, and thus activities of “DA” neurons in response to US presentation are diminished. As a result, no enhancement of synapses from “CS2” neurons to “CR” neurons occurs by subsequent compound conditioning of CS1 and CS2 with the US, and thus no learning of CS2 occurs. Synapses for which efficacy can be changed by conditioning are colored in red and marked as “modifiable”. Excitatory synapses are marked as triangles; inhibitory synapses are marked as bars. UR: unconditioned response.

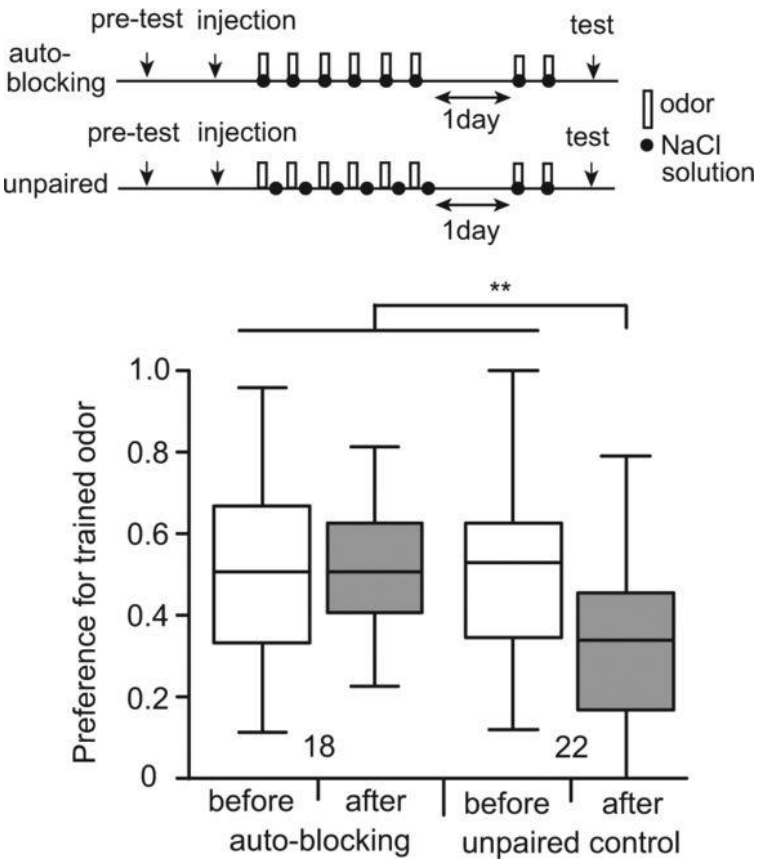


Figure 5. Auto-blocking. Two groups of animals received a pre-test and were then injected with 3 μl of saline containing 100 μM flupentixol. Thirty min later, one group (auto-blocking group) was subjected to 6-trial pairing of an odor with NaCl solution and the other group (unpaired control group) was subjected to unpaired presentation of an odor and NaCl solution 6 times each. The ITI was 5 min for the former and it was 2.5 min for the latter. On the next day, both groups were subjected to 2-trial pairing of the odor with NaCl solution and 20 min later they received a post-test. Relative odor preferences for the trained odor before (while boxes) and after (gray boxes) training are shown as box and whisker diagrams. The number of animals is shown below the boxes. A GLMM was used for comparison of relative preferences for the trained odor before and after conditioning in the auto-blocking and control groups (Supplemental table S1). Statistical significance is shown as asterisks (** $p<0.01$).

Table 1. Procedures used for and results of compound conditioning, blocking and auto-blocking experiments.

Group	Phase 1	Phase 2	Results: Learning of Y?	Figures
Compound	-	XY+	Yes	Fig. 2
Control	-	Y+	Yes	Fig. 2
Blocking	X+	XY+	No	Fig. 3
Unpaired control	X / +	XY+	Yes	Fig. 3
Auto-blocking	Y+ (under flupentixol)	Y+	No	Fig. 5
Control	Y / + (under flupentixol)	Y+	Yes	Fig. 5
Auto-blocking	Y+ (under epinastine)	Y+	Yes	Fig. S1

XY+: a compound of stimulus X and stimulus Y is paired with aversive US; Y / +: unpaired presentation of stimulus Y and aversive US.

Supplemental Figures and Tables

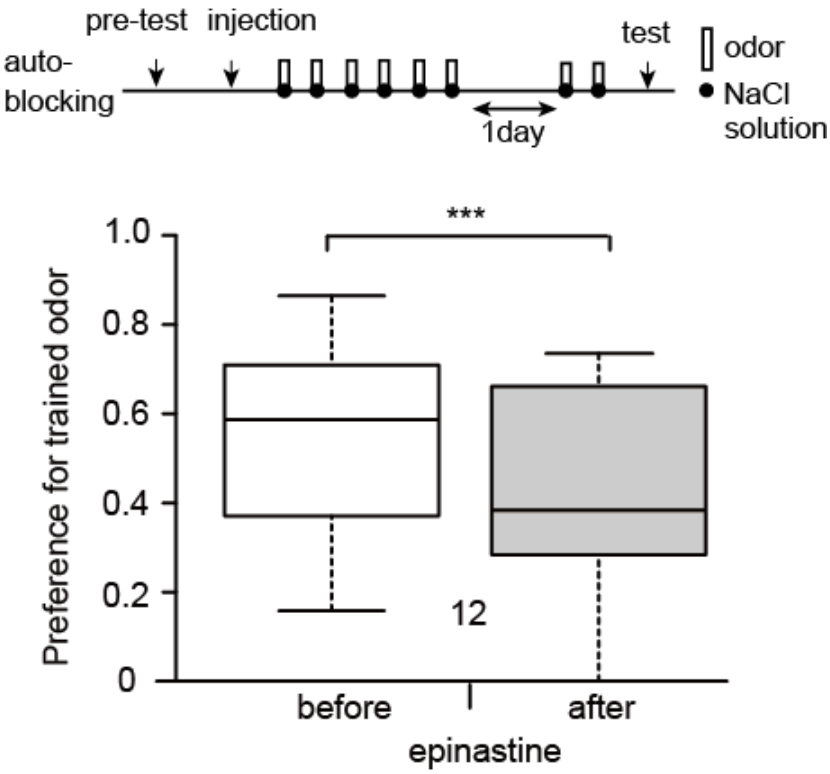


Figure S1. Auto-blocking does not occur by epinastine. One group of animals received a pre-test and was then injected with 3 μl of saline containing 2 μM epinastine. Thirty min later, they were subjected to 6-trial pairing of an odor with NaCl solution. The ITI was 5 min for. On the next day, they were subjected to 2-trial pairing of the odor with NaCl solution and 20 min later they received a post-test. Relative odor preferences for the trained odor before (while boxes) and after (gray boxes) training are shown as box and whisker diagrams. The number of animals is shown below the boxes. A GLMM was used for comparison of relative preferences for the trained odor before and after conditioning (Supplemental table S1). Statistical significance is shown as asterisks (*** $p<0.001$).

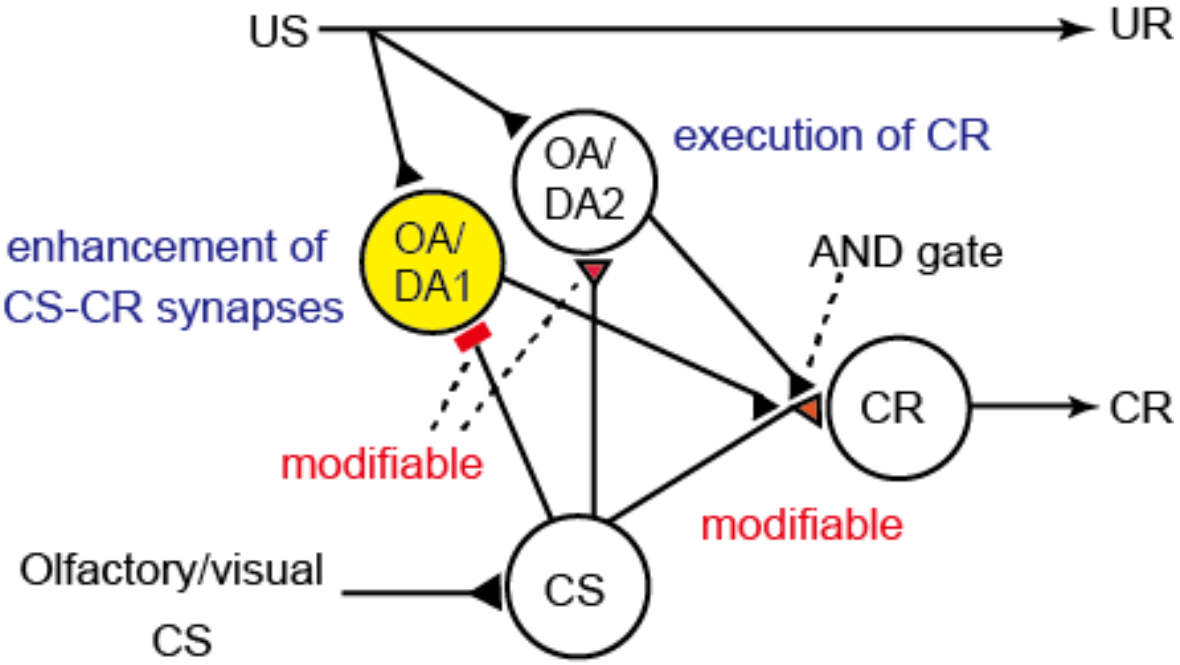


Figure S2. Complete description of my model of appetitive and aversive learning in crickets. The model was proposed to account for blocking and auto-blocking of appetitive learning[11] and aversive learning (this study), by modifying a previous model[15] that is proposed to account for findings that blockade of OA- or DA-ergic transmission impairs learning and execution of conditioned response (or memory retrieval) in appetitive or aversive learning, respectively, in crickets[12-19]. The model assumes two classes of OA and DA neurons, namely, the OA1 and DA1 neurons ("OA/DA1" neurons; colored in yellow) that govern enhancement of "CS-CR" synapses (but not execution of CR) and "OA/DA2" neurons that govern execution of CR or memory retrieval (but not enhancement of "CS-CR" synapses). The model also assumes that (1) "CS" neurons (which may represent intrinsic neurons of the mushroom body) that convey signals about CS make silent or weak synaptic connections with dendrites of "CR" neurons (which may represent efferent (output) neurons of the mushroom body lobe), activation of which leads to a conditioned response (CR), but these synaptic connections are silent or very weak before conditioning, (2) The "OA/DA1" neurons are assumed to receive excitatory synapses that represent US signal and silent or very weak inhibitory synapses from "CS" neurons before training, which are strengthened by CS-US pairing. (3) During training, "OA/DA1" neurons receive excitatory synaptic input that represents actual US and inhibitory input from "CS" neurons that

represents US prediction by CS, and thus their activities represent US prediction error signals (see legends of Fig. 4). (4) The “OA/DA2” neurons are assumed to receive excitatory synapses that represent US signal and silent or very weak excitatory synapses from “CS” neurons before training, which are strengthened by CS-US pairing. (5) “OA/DA2” neurons make synaptic connections with axon terminals of “CS” neurons, and coincident activation of “CS” neurons and “OA/DA2” neurons is needed for activation of “CR” neurons (AND gate) and for production of conditioned response. Presentation of CS after CS-US pairing activates “CS” neurons and then “OA/DA2” neurons and thus activates “CR” neurons to lead to conditioned response. UR: unconditioned response.

Supplemental table S1. Summary of generalized linear mixed model (GLMM).

a. Compound conditioning experiment for olfactory learning (Fig. 2a)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	0.2939	0.1913	1.537	0.124
Test	-0.9557	0.1559	-6.132	$8.68 * 10^{-10}$
Training	-0.08869	0.2446	-0.363	0.717
Test * Training	0.2947	0.2004	1.471	0.141

b. Compound conditioning experiment for visual learning (Fig. 2b)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	0.1952	0.2809	0.695	0.487
Test	-1.2552	0.1616	-7.768	$7.97 *10^{-15}$
Training	-0.4594	0.3954	-1.162	0.245
Test * Training	0.7840	0.2053	3.818	$1.35 *10^{-4}$

c. Blocking experiment for olfactory learning (Fig. 3a)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	0.1639	0.1398	1.173	0.241
Test	-0.1943	0.1455	-1.336	0.182
Training	0.06987	0.1930	0.362	0.717
Test * Training	-0.6110	0.1922	-3.179	0.00148

d. Blocking experiment for visual learning (Fig. 3b)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	-0.2931	0.2954	-0.992	0.321

Test	0.1979	0.1573	1.258	0.208
Training	-0.1123	0.4255	-0.264	0.792
Test * Training	-1.330	0.2414	-5.509	$3.6 * 10^{-8}$

e. Auto-blocking experiment with flupentixol in olfactory learning (Fig. 5)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	0.03878	0.1618	0.240	0.811
Test	-0.01699	0.1348	-0.126	0.900
Training	-0.007372	0.2152	-0.034	0.973
Test * Training	0.5831	0.1830	-3.186	0.00144

f. Auto-blocking experiment with epinastine in olfactory learning (Supplemental Fig. S1)

Fixed effects	Estimate	Standard error	Z value	P value
Intercept	0.2268	0.2314	0.980	0.327
Training	-0.5928	0.1753	-3.381	$7.23 * 10^{-4}$

By using a GLMM, effects of the test situation (before or after training), training procedure (compound conditioning or control procedure in a and b, blocking or control procedure in c and d, auto-blocking or control procedure in e) and interaction between the test and training on relative preference for the trained odor or pattern were evaluated. The estimate for the intercept indicates the estimate before training in the compound group (a and b), blocking group (c and d) or auto-blocking group (e and f).

Supplemental table S2. Information coded in the responses of DA neurons in the aversive learning model.

Stimulus	Before training	After training
US	1(US)	1 (US)
CS	0	0 [-1 (-USP)]*
CS+US	1 (US)	0 (USPE)

The table shows responses of DA neurons in the model shown in Fig. 4a to aversive US, CS and paired presentation of the CS and US before and after training. DA neurons govern enhancement of synaptic transmission that underlies conditioning (CS-CR synapse). After completion of training, paired presentation of CS and US does not produce responses in DA neurons and thus no further enhancement of synaptic transmission occurs. USP: US prediction; USPE: US prediction error. Responses are indicated as all or none (1 or 0). *Negative value in parentheses indicates inhibitory synaptic input.

General Discussion

In chapters 1 and 2, I provided evidence of a blocking phenomenon in crickets. Then I proposed a new learning model that matches the prediction error-based learning theories and evaluated the model by an auto-blocking experiment. The results provide rigorous evidence of prediction error-based learning in classical conditioning. The results also suggested that octopamine (OA) and dopamine (DA) convey reward and aversive prediction error signals, respectively. Moreover, I provided distinctive evidence for the first time that supports the Rescorla-Wagner model by demonstrating one-trial blocking and auto-blocking in appetitive learning. Here I discuss some general questions to help extension of the study.

How are the negative prediction error signals coded by aminergic neurons?

An important experiment to be performed is an extinction experiment. According to prediction error-based learning theories, association is reduced when the actual US of the conditioning trial is less than expected, in which negative prediction error is produced. An extinction experiment is used to examine negative prediction error. After associative training for a CS paired with US, an associated CS is presented in the absence of the US. This generates a negative prediction error to promote a decrement of conditioned response (CR).

Neurons that mediate negative predictor error signals for aversive learning remain to be identified. In mammals, it is well accepted that DA neurons encode prediction error during appetitive conditioning[1, 2]. In the context of extinction, omission of an expected reward causes depressions of the DA neurons' firing at the time that the reward was expected. These depressions are discussed as reward negative prediction error signals. On the other hand, evidence for negative prediction error signals in aversive learning is limited[3]. It has been suggested that DA neurons encode value prediction error signals including positive aversive US prediction error[4]. However, it remains unclear whether

they encode negative aversive US prediction error signals. Lateral amygdala (LA) neurons are activated for unpredicted aversive US, and such a response is decreased across the course of CS-US pairing in fear conditioning[5]. It seems that LA neurons convey positive aversive US prediction error signals, but there has been no report showing that they convey negative aversive US prediction error signals in an extinction trial.

In crickets, it remains to be determined whether OA or DA neurons convey appetitive or aversive negative prediction error signals. In relation to this subject, it should be noted that my present model does not account for extinction. However, it is feasible to modify the model to account for extinction by assuming that OA or DA neurons have spontaneous activities so that they can convey positive prediction error signals as excitation and negative prediction error signals as inhibition. Excitation of OA or DA neurons strengthens CS-CR synapses and inhibition of OA or DA neurons weakens the synapses. Evaluation of this possibility remains as a future subject.

Do the prediction error-based learning theories represent a general learning principle beyond different taxa?

My study demonstrated prediction error-based learning in insects, animals that are evolutionarily remote from mammals. An important research subject is to examine how ubiquitous prediction error-based learning is among animal phyla. To discuss this subject, I summarize previous reports on classical conditioning and blocking phenomena in different taxa.

Reports on classical conditioning in various animals and other organisms are summarized in a phylogenic tree in Fig. 1, on the basis of the available literature[6-14]. Classical conditioning has been reported in most animal taxa and even in the chromista[13] and plantae[14]. It would be interesting to know whether associative learning capability evolved independently in different taxa or whether it is derived from common ancestors.

Blocking has been reported in platyhelminths, mollusks, arthropods, and vertebrates.

Researchers studying platyhelminths or mollusks prefer the term “cue competition” to account for blocking[11, 12], but experiments to discriminate alternative theories have not been performed. Whether blocking in platyhelminthes and mollusks can be accounted for by prediction error-based learning is a critical issue to clarify to what extent classical conditioning based on prediction error is ubiquitous among animal phyla.

Finally, I discuss the possible adaptive significance of associative learning based on prediction error. One of the benefits would be to extract sensory cues to predict biologically significant events in complex environmental conditions, where many stimuli exist. Consider a case in which stimulus X is always presented with a US, and animals experience X and a US in various environmental conditions where stimulus A, B or C co-exists; in other words, they experience X+, AX+, BX+, and CX+ conditioning. If animals associate each of the cues with the US on the basis of temporal contiguity, they will become responsive to all of the A, B, C and X signals. On the other hand, associative learning based on prediction error is more economical, allowing learning to respond to X and inhibiting learning for the less important signals A, B and C. Thus, it is evident that learning based on prediction error is more beneficial for survival of animals than that based on temporal contiguity in a complex environment. However, is the cost of maintaining a neural mechanism to achieve such an advanced form of learning too high? In mammals, costly, redundant and complex mechanisms have been suggested[15]; information for reward prediction error computations is widely distributed in many brain regions and some kinds of characteristics are already mixed in input neurons of the dopamine neurons. However, my model suggests that calculation of prediction error can be achieved even by a simple neural circuit. Thus, the cost of maintaining this circuit may not be so high. I suggest that associative learning based on prediction error will be found in many taxa, including mollusca and planarians, in the future. Auto-blocking experiments should help to discriminate prediction error-based learning from alternative possible learning mechanisms if neurotransmitters conveying US signals are found in

these animals.

References

1. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science.* **275**, 1593–1599 (1997).
2. Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57**, 87–115 (2006).
3. Li, S. S. Y. & McNally, G. P. The conditions that promote fear learning: Prediction error and Pavlovian fear conditioning. *Neurobiol. Learn. Mem.* **108**, 14–21 (2014).
4. Matsumoto, H., Tian, J., Uchida, N. & Watabe-Uchida, M. Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *Elife.* **5**, 1–24 (2016).
5. Johansen, J. P., Tarpley, J. W., LeDoux, J. E. & Blair, H. T. Neural substrates for expectation-modulated fear learning in the amygdala and periaqueductal gray. *Nat. Neurosci.* **13**, 979–986 (2010).
6. Perry, C. J., Barron, A. B. & Cheng, K. Invertebrate learning and cognition: Relating phenomena to neural substrate. *Wiley Interdiscip. Rev. Cogn. Sci.* **4**, 561–582 (2013).
7. Mizunami, M., Yokohari, F. & Takahata, M. Exploration into the adaptive design of the arthropod 'Microbrain'. *Zool. Sci.* **16**, 703–709 (1999).
8. Byrne, J. H. & Hawkins, R. D. Nonassociative Learning in Invertebrates. *Cold Spring Harb Perspect Biol.* **7**, 1–16 (2015). doi:10.1101/cshperspect.a021675
9. Kodama, E. et al. Insulin-like signaling and the neural circuit for integrative behavior in C. elegans. *Genes. Dev.* **20**, 2955–2960 (2006). doi:10.1101/gad.1479906
10. Mohri, A. et al. Genetic control of temperature preference in the nematode *Caenorhabditis elegans*. *Genetics* **169**, 1437–1450 (2005).
11. Prados, J. et al. Blocking in rats, humans and snails using a within-subjects design. *Behav. Processes*. **100**, 23–31 (2013).
12. Prados, J. et al. Cue competition effects in the planarian. *Anim. Cogn.* **16**, 177–186 (2013).
13. Hennessey, T. M., Rucker, W. B. & McDiarmid, C. G. Classical conditioning in

paramecia. *Anim. Learn. Behav.* **7**, 417–423 (1979).

14. Gagliano, M., Vyazovskiy, V. V., Borbély, A. A., Grimonprez, M. & Depczynski, M. Learning by Association in Plants. Sci. Rep. **6**, 1–9 (2016).
15. Tian, J. *et al.* Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons. *Neuron.* **91,** 1374–1389 (2016).
16. Giribet, G. New animal phylogeny: future challenges for animal phylogeny in the age of phylogenomics. *Org. Divers. Evol.* **16,** 419–426 (2016).

Figures

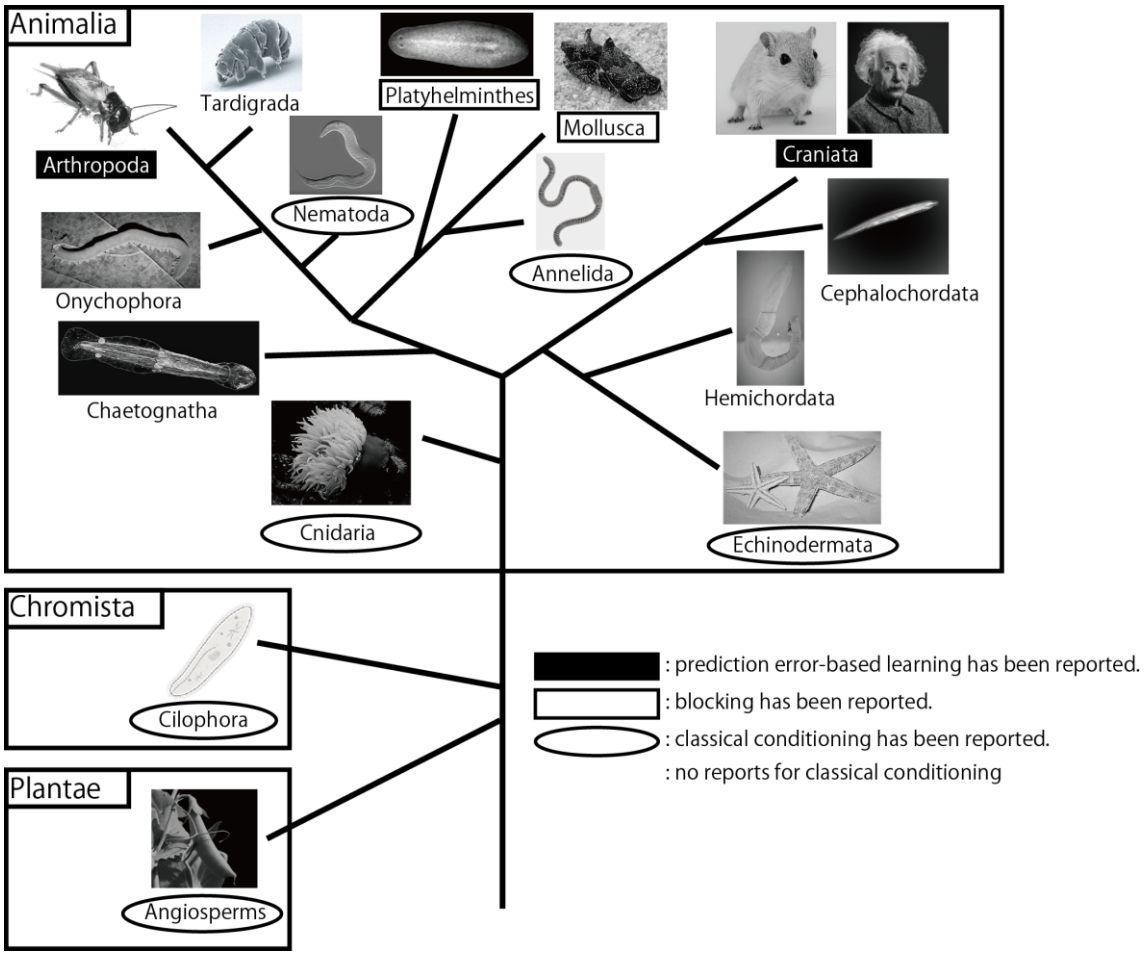


Fig. 1 Learning ability in various phyla

Reports of learning ability are summarized in a phylogenic tree. The dendrogram shown is based on Giribet (2016)[16]. Photos are originated from following websites: https://pixabay.com, https://en.wikipedia.org, or http://www.kenq.net/ill/. They are free to use or licensed by Creative Commons (https://creativecommons.org/licenses).

Achievement

List of publications

1) Terao K., Matsumoto Y., Mizunami M.

『Critical evidence for the prediction error theory in associative learning』

Scientific Reports, Nature Publication Group, *5*, 2015（peer reviewed）

2) Terao K., Mizunami M.

『Roles of dopamine neurons in mediating the prediction error in aversive learning in insects』

Scientific Reports, Nature Publication Group, 7, 2017（peer reviewed）

List of other publications

(1) Manuscripts

1) Matsumoto Y., Hirashima D., Terao K., Mizunami M.
『**Roles of NO signaling in long-term memory formation in visual learning in an insect**』
PLOS ONE, Public Library of Science, 8(7), 2013: e68538（peer reviewed）

2) Mizunami M., Nemoto Y., Terao K., Hamanaka Y., Matsumoto Y.
『**Roles of calcium/calmodulin-dependent kinase II in long-term memory formation in crickets**』
PLOS ONE, Public Library of Science, 9(9), 2014: e107442 （peer reviewed）

3) Awata H., Wakuda R., Ishimaru Y., Matsuoka Y., Terao K., Katata S., Matsumoto Y., Hamanaka Y., Noji S., Mito T., Mizunami M.
『**Roles of OA1 octopamine receptor and Dop1 dopamine receptor in mediating appetitive and aversive reinforcement revealed by RNAi studies**』
Scientific Reports, Nature Publication Group, 6, 2016（peer reviewed）

4) Mizunami M, Hirohata S, Sato S, Arai R, Terao K, Sato M, Matsumoto Y
『**Habit formation by extended pavlovian training**』
under review

5) Hasegawa E*, Aonuma H*, Terao K*, Ogusu N, Ohkubo Y, Watanabe S, Mikami S, Fujita Y, Mizunami M, Murakami Y (* equally contributed)
『**Contact with nest mates maintains a self-sacrificing motivation for social defenses in an ant**』
under review

(2) Presentations in International Conferences

1) ○Terao K., Matsumoto Y., Mizunami M.,
『**Critical evidence for the prediction error theory in associative learning**』
Hokkaido Neuroethology Workshop 2014; 4. Small brains, bright minds: Learning and memory in invertebrates, 8, Sapporo, Japan,
July, 2014（poster presentation）

2) ○Terao K., Matsumoto Y., Mizunami M.,

『**Critical evidence for the prediction error theory in associative learning**』
A joint meeting of the 11th International Congress of Neuroethology and the 36th Annual Meeting of the Japanese Society for Comparative Physiology and Biochemistry, PO-1097, Sapporo, Japan,
July, 2014（poster presentation）

3) Ewen-Campen B., Wakuda R., Terao K., Matsumoto Y., Mizunami M., Extavour C.,
『***oskar* functions in adult neural stem cells to influence long-term memory formation in the cricket *Gryllus bimaculatus***』
A joint meeting of the 11th International Congress of Neuroethology and the 36th Annual Meeting of the Japanese Society for Comparative Physiology and Biochemistry, PO-2100, Sapporo, Japan,
July, 2014（poster presentation）

4) Lopez DH, Ewen-Campen B., Wakuda R., Terao K., Matsumoto Y., Mizunami M., Extavour C.,
『**Functional Characterization of the Role of the Pole Plasm Component Oskar in the Adult Brain of the Cricket *Gryllus bimaculatus***』
Society of Developmental Biology 74th Annual Meeting, Development and Evolution, 206, B12, Snowbird, Utah, USA
July, 2015（poster presentation）

5) ○Terao K., Matsumoto Y., Mizunami M.
『**Crucial evidence of the prediction error theory in a behavioral pharmacological study** 』
Neuroscience 2015 Society for Neuroscience Annual Meeting, Invertebrate Learning and Memory II, 630.05, Chicago, Illinois, USA
October, 2015（poster presentation）

6) ○Kanta TERAO,
『**Critical evidence for the prediction error theory in the insect learning**』
Avian Brain and Behaviour, Sapporo Workshop 2016 “Learning and cognition embedded in animal behaviours”, Hokkaido University
November, 2016（oral presentation）

7) ○Kanta TERAO, Yukihisa MATSUMOTO, Makoto MIZUNAMI
『**Prediction error theory in insects; blocking experiment and pharmacological evaluation**』
The 44th Naito Conference “Decision Making in the Brain—Motivation, Prediction, and Learning”, Sapporo, PS[II]-12,
October, 2017（poster presentation, reviewed）

(3) Presentations in Domestic Conferences

1) ○寺尾 勘太，松本 幸久，水波 誠

『コオロギにおけるブロッキング現象の確認』

日本動物学会北海道支部第58回大会, O-1, 北海道教育大学, 2013年8月（口頭発表・査読無し）

2) ○寺尾 勘太，松本 幸久，水波 誠

『コオロギの古典的条件づけにおけるブロッキング』

日本動物学会第84回岡山大会, 1K1415, 岡山大学, 2013年9月（口頭発表・査読無し）

3) ○寺尾 勘太，松本 幸久，水波 誠

『予測誤差に基づく昆虫の学習』

日本動物学会第85回仙台大会, 3E1400, 東北大学, 2014年9月（口頭発表・査読無し）

4) ○寺尾 勘太，水波 誠

『ドーパミンニューロンの伝達する罰予測誤差』

日本動物学会第86回新潟大会, 1D1715, 新潟コンベンションセンター 朱鷺メッセ, 2015年9月（口頭発表・査読無し）

5) ○寺尾 勘太，松本 幸久，水波 誠

『コオロギは『ビックリ』することで学習する』

日本動物行動学会第34回東京大会, P098, 東京海洋大学, 2015年11月（ポスター発表・査読無し）

6) ○寺尾 勘太，水波 誠

『Aminergic neurons convey information of prediction error in crickets』

第40回日本比較内分泌学会・第37回日本比較生理生化学会合同大会（CompBiol 2015広島大会）, PO-1, JMS アステールプラザ, 2015年12月（口頭発表およびポスター発表・査読無し）

7) ○Kanta TERAO, Takayuki WATANABE, Hitoshi AONUMA, Hiroshi NISHINO, Makoto MIZUNAMI

『Distribution of putative octopaminergic and tyraminergic neurons in a cockroach brain』

第38回日本比較生理生化学会東京大会, P-018, 玉川大学, 2016年9月（ポスター発表・査読無し）

8) ○Kanta TERAO, Makoto MIZUNAMI

『Critical evidence for the prediction error theory in the insect learning』

日本動物心理学会第76回北海道大会, O-24-06, 北海道大学, 2016年11月（口頭発表・査読無し）

9) 長谷川英祐, 青沼仁志, ○寺尾勘太, 小楠なつき, 大久保祐作, 渡邊紗織, 三上俊太, 藤田悠介, 水波誠, 村上優花
『アリは巣仲間と交流してストレスから回復する』
行動2017 日本動物関連学会・研究会　合同大会, P-185, 東京大学, 2017年8月（ポスター発表・査読無し）

10) ○Kanta TERAO, Makoto MIZUNAMI
『コオロギの罰学習におけるブロッキング現象』
日本動物学会第88回富山大会, 1G0915, 富山県民会館, 2017年9月（口頭発表・査読無し）

11) ○Kanta TERAO, Makoto MIZUNAMI
『Prediction error theory in insects; comparative analyses produce insights for general learning rules』
日本基礎心理学会第36回大阪大会, 若手オーラルセッション ファイナリスト, 立命館大学, 2017年12月（口頭発表予定・査読有り）

Acknowledgement

I would like to express my sincere gratitude to Dr. Makoto Mizunami for his kind and helpful supports as my supervisor. I also thank Dr. Yukihisa Matsumoto for his technical guidance and assistance.

I appreciate Drs. Toshiya Matsushima (Hokkaido University) and Yutaka Kosaki (Waseda University) for their helpful review comments and suggestions.

I'd like to thank all teachers and colleagues in department of animal behavior and system control. I also thank my friends YM, TK, TY, KM, and NN for their support along these years.

Finally, I would like to thank my family.

My studies are supported by fellowship for young scientists (Japan Society for Promotion of Science, 15J01414).