A neuromorphic model of active vision shows spatio-temporal encoding in lobula neurons can aid pattern recognition in bees

HaDi MaBouDi[1,2,3*♠], Mark Roper[3,4], Marie-Genevieve Guiraud[3,5], Lars Chittka[3], James A.R. Marshall[1,2♠]

[1]Department of Computer Science, University of Sheffield, Sheffield, UK
[2]Neuroscience Institute, University of Sheffield, Sheffield, UK
[3]School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
[4]Drone Development Lab, Ben Thorns Ltd, Colchester, Essex, UK
[5]School of Natural Sciences, Macquarie University, North Ryde, NSW, Australia

*Corresponding author: HaDi MaBouDi, maboudi@gmail.com
♠ Current address: Opteran Technologies, Sheffield Innovation Centre, Sheffield, UK

Short title: Insect-inspired minimal model for active vision

Abstract

Bees possess remarkable cognitive abilities in on-the-fly visual learning, making them an ideal model for studying active information acquisition and representation. In this study, we investigated the minimal circuitry required for active vision in bees by considering their flight behaviours during visual pattern scanning. By developing a neural network model inspired by the insect visual system, we examined the influence of scanning behaviour on optic lobe connectivity and neural activity. Through the incorporation of non-associative learning and exposure to diverse natural images, we obtained compelling results that align with neurobiological observations. Our findings reveal that active scanning and non-associative learning dynamically shape the connectivity within the visual lobe, resulting in an efficient representation of visual input. Interestingly, we observed self-organization in orientation-selective neurons in the lobula region, characterized by sparse responses to orthogonal bar movements. These dynamic orientation-selective cells cover various orientations, exhibiting a bias towards the speed and contrast of input sampling. To assess the effectiveness of this spatiotemporal coding for pattern recognition, we integrated our model with the mushroom body circuitry underlying associative learning. Notably, our model demonstrated impressive performance across several pattern recognition tasks, suggesting a similar coding system within the bee visual system. Overall, this study integrates behavioural experiments, neurobiological findings, and computational models to reveal how complex visual features can be condensed through spatiotemporal encoding in the lobula neurons, facilitating efficient sampling of visual cues for identifying rewarding foraging resources. Our findings have broader implications for understanding active vision in diverse animals, including humans, and offer valuable insights for the application of bio-inspired principles in the design of autonomous robots.

Keywords

active vision, image statistics, lobula, mushroom bodies, non-associative learning, orientation selective neurons, scanning behaviour, visual recognition.

Introduction

Bees are capable of remarkable cognitive feats, in particularly in visual learning (Srinivasan, 2010; Turner, 1911; Von Frisch, 1914; Wehner, 1967); they can not only learn to associate a colour or orientation of a bar with reward (Dyer et al., 2011; Srinivasan, 1994; Stach et al., 2004) but are also able to identify specific features to categorise visual patterns, by finding the relevant stimuli properties (Benard et al., 2006; Stach et al., 2004). Furthermore, bees have demonstrated the capacity to grasp abstract concepts (Avarguès-Weber et al., 2011; Giurfa et al., 2001; Guiraud et al., 2018; MaBouDi et al., 2020c; Menzel, 2012) and solve numerosity tasks by scanning the elements within the presented stimuli (MaBouDi et al., 2020a). These exceptional capabilities position bees as a valuable animal model for investigating the principles of visual learning through the analysis of their behavioural responses (Menzel and Giurfa, 2006; Srinivasan, 2010). Nevertheless, it is still unclear how bees, with low visual acuity (Gribakin, 1975; Srinivasan and Lehrer, 1988) and limited neural resources can recognise complex patterns, and indeed perceive the visually intricate natural world they encounter during daily foraging activities (Chittka and Niven, 2009; Giurfa, 2013).

Bees, as vital pollinators, must cope with the variability of the natural environment for their survival. The natural scenes that animals typically encounter are structured differently to random/non-natural scenes (Matthews et al., 2018; Ruderman, 1994; Simoncelli and Olshausen, 2001; Zimmermann et al., 2018). It has been hypothesised that visual sensory neurons must efficiently adapt to the regularities of the natural scene to take advantage of this spatio-temporal structure, and to efficiently code the information in the visual environment (Barlow, 1961). Hence, Insects visual neurons have evolved to provide robust and efficient responses to naturalistic inputs, enabling survival in complex ecological niches (Dyakova et al., 2019, 2015; Dyakova and Nordström, 2017). Numerous studies have highlighted the remarkable adaptability of the sensory pathway in insects, demonstrating robust response and behaviour across various input parameters, such as contrast, spatial frequency and spatiotemporal correlations (Arenz et al., 2017; Brinkworth and O'Carroll, 2009; Clark et al., 2014; Dyakova et al., 2019; Dyakova and Nordström, 2017; Juusola and Song, 2017; Schwegmann et al., 2014; Serbe et al., 2016; Song and Juusola, 2014, 2014; Van Hateren, 1997; van Hateren, 1992). For instance, Song and Juusola (2014) showed that fly photoreceptors extract more information from naturalistic time series compared to artificial stimuli or white noise, exhibiting stranger responses with a higher signal-to-noise ratio (Song and Juusola, 2014). Despite insights gained from the statistical properties of the natural scene, the precise neural mechanisms underlying visual processing of natural scene remain elusive, necessitating further investigation. This study aims to investigate the theoretical aspects of how insect visual circuitry adapts to regularities in natural scene. One of our objectives is to understand the efficient coding strategies and robust response mechanisms employed by these neurons that play a crucial role in enhancing visual pattern recognition and facilitate survival and navigation in variable environments.

In the realm of animal vision, active vision strategies have been observed, where animals actively scan tagets and extract visual information over time (Land, 1999; Land and Nilsson, 2012; Yarbus, 2013). Primates employ eye movements, including microsaccades, as an effective sampling strategy that enhances fine spatial information and improves the encoding details of natural stimuli (Anderson et al., 2020; Land, 1999; Rucci et al., 2007; Rucci and Victor, 2015). Other species, such as insects, exhibit active vision strategies throught characteristic head/body-movements or a certain approaching paths during a visual tasks (Chittka and Skorupski, 2017; Dawkins and Woodington, 2000; Land, 1973; Land and Nilsson, 2012; Langridge et al., 2021). Recent studies have discovered that Drosophila are able to move their retinas to stabilise their retinal images, achieving hyperacute vision and enhancing depth perception (Fenk et al., 2022; Juusola et al., 2017). Honeybees may need to sample and integrate colour information due to their limited ability to discriminate similar colours in brief flashes (<50ms) (Nityananda et al., 2014). They exhibit a sequence of movements in response to particular visual stimuli and employ continuous sampling to build a representation of their environment (Boeddeker et al., 2015; Collett et al., 1993; Doussot et al., 2021; Guiraud et al., 2018; Langridge et al., 2018; Lehrer and Collett, 1994; MaBouDi et al., 2020a; Werner et al., 2016). For instance, bumblebees acquire visual details of numerosity by sequential scanning of stimulus elements rather than parallel processing to enumerate the countable elements within the visual displays (MaBouDi et al., 2020a). A recent analysis of bumblebees' flight trajectories showed that they sequentially scanned specific regions of the patterns prior to making a decision, instead of global pattern processing (Langridge et al., 2021; MaBouDi et al., 2021b). Therefore, the low-resolution compound eyes and the possibly reduced parallel processing in insects (compared to vertebrates) suggest that these bees may employ an active vision strategy by continuous sampling to build up a picture of their environment (Chittka and Skorupski, 2017; Nityananda et al., 2014). These active vision strategies, akin to primate eye movements, serve fundamental functions in early visual processing for redundancy reduction (Doussot et al., 2021; Kuang et al., 2012; Odenthal et al., 2021). Nonetheless, it is still poorly understood how active vision strategies enable bees to overcome their limited representational capacity, discover regularities in the visual input and solve complex visual discriminations.

Taking advantage of the bee's visual ability and the detailed information on flight's paths analysed in our previous study when bees solved a simple visual task (MaBouDi et al., 2021b), we conducted further research to investigate the necessary and minimally sufficient circuitry required for active vision of achromatic pattern recognition in bees. Importantly, our primary objective was to investigate how the scanning behaviour of bees contributes to the functional organization and connectivity of neurons in the visual lobe. We specifically focused on the hypothesis that complex visual features can be efficiently encoded through spatio-temporal patterns of activity in the lobula neurons, leading to distinct and specific representations necessary for learning in the miniature brain of bees. Through our study, we aimed to provide valuable insights into the intricate interplay between scanning behaviour and neural activity within the bee visual system, ultimately advancing

our understanding of the mechanisms underlying active vision. To achieve this, we developed a neuromorphic model of the bee optic lobes, incorporating the concept of efficient coding through the implementation of non-associative plasticity. Through this mode, we demonstrated how spatial scanning behaviour of bees in response to naturalistic visual inputs shape the connectivity within the medulla (2nd optic ganglion) and facilitate an efficient representation of these inputs in the lobula (3rd optic ganglion). This efficiency is achieved through the self-organization of a specific set of orientation-selective neuron in the lobula, highlighting the combined impact of scanning behaviour and non-associative learning on shaping the neural circuitry within the bees' optic lobes. To assess the model, we integrated the optic lobe model with a secondary decision-making neural network inspired by neural mechanisms of associative learning in insect brains and supported by previous neurobiological findings (Cassenaer and Laurent, 2012; Okada et al., 2007; Paulk et al., 2009; Paulk and Gronenberg, 2008; Yang and Maddess, 1997). Visual input flight dynamics for the model were derived from our observations of bee behaviour during a visual discrimination task (MaBouDi et al., 2021b). This allowed us to evaluate and test the hypothesis of active sampling from our model against real-world behaviour results (MaBouDi et al., 2021b), as well as other published visual discrimination tasks performed by bees (Benard et al., 2006; Dyer et al., 2005; Srinivasan, 2010, 1994; Zhang and Horridge, 1992). Furthermore, we conducted a detailed analysis comparing the features and properties of neural responses that emerged in our model with existing neurobiological findings (James and Osorio, 1996; Paulk et al., 2008; Seelig and Jayaraman, 2013; Yang and Maddess, 1997). By aligning our model's outputs with established neurobiological data, we enhanced the credibility and reliability of our model in accurately capturing essential aspects of neural processing associated with active vision.

Results

A bio-inspired neural network of active vision

To understand how bee scanning behaviour may efficiently shape the activity of neurons in the visual lobes of the bee brain based on the efficient coding, and how the visual information is processed for pattern recognition, we designed a neural network that drew inspiration from the known morphological and functional features of the insect brain (Figure 1A, C). The network approximates the neural circuitry which initially processes the visual input within the bees' lamina and medulla (1st an 2nd optic ganglia). Then, to replicate temporal encoding during scans (Figure 1B), we supposed a structure of time delay of between 1 - 5 'temporal instances' from the medulla neuron outputs to the lobula (3rd optic ganglion) wide-field neurons (Figure 1D). This arrangement enabled sequential sampling of specific locations along the scan line of the viewed visual

137 pattern, resulting in a gradually accumulated internal representation of the visual input as the ultimate output
138 of the lobula neurons.
139

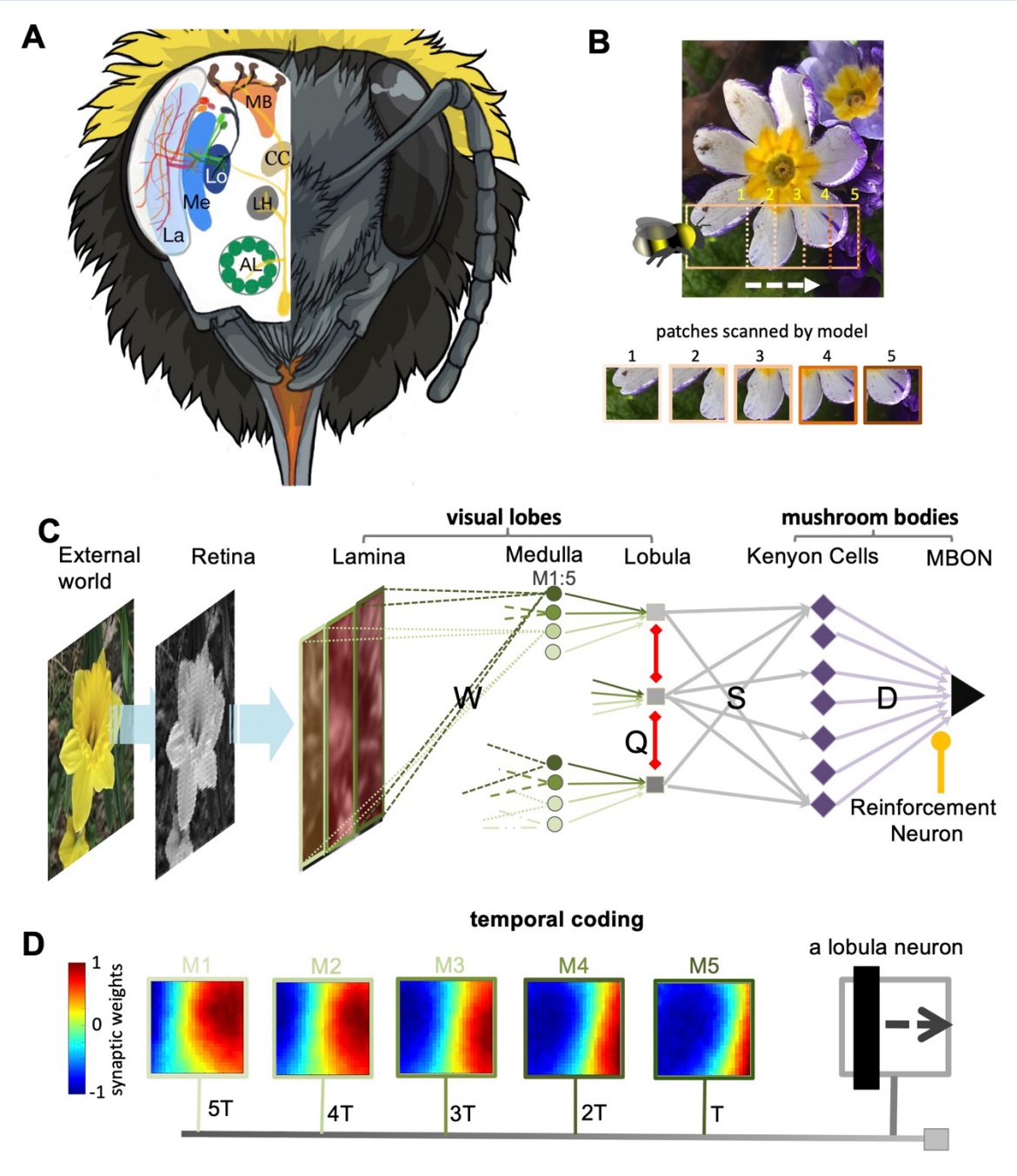


Figure 1. Neural network of active vision inspired by neurobiology and flight dynamics of bees. (A) The right side displays the front view of the bumblebee head showing the component eye and antenna. Left hand side presents a schematic view of the bee's brain regions. Part of neural pathways from the retina to the mushroom bodies are also represented. Labels: AL – Antennal lobe; LH – Lateral horn; CC – Central complex; La – Lamina; Me – Medulla; Lo – Lobula; MB – Mushroom body. Figure was designed by Alice Bridges **(B)** A representation of the modelled bee's scanning behaviour of a flower demonstrating how a sequence of patches project to the simulated bee's eye with lateral movement from left to right. Below

shows the five image patches sampled by the simulated bee. **(C)** Representation of the neural network model of active vision inspired by micromorphology of the bee brain that underlie learning, memory and experience-dependent control of behaviour. The photoreceptors located in the eye component are excited by the input pattern. The activities of photoreceptors change the membrane potential of a neuron in the next layer, Lamina. The lamina neurons send signals (through *W* connectivity matrix) to the medulla neurons to generate spikes in this layer. Each wide-field lobula neuron integrates the synaptic output of five small-field medulla neurons. The lobula neurons are laterally inhibited by local lobula interconnections (via *Q* connectivity matrix). Lobula neurons send axons into the mushroom body for connection with Kenyon Cells (KCs) through a random matrix of connectivity, *S*. The KCs all connect to a single mushroom body output neuron (MBON) through random synaptic connections *D*. A single reinforcement neuron (yellow neuron) modulates the synaptic weights between KCs and MBON by simulating the release of octopamine or dopamine when presented with specific visual stimuli (see Method section). **(D)** A temporal coding model that is proposed as the connectivity between medulla and lobula neurons. Each matrix shows the inhibitory (blue) and excitatory (red) connectivity between lamina neurons to a medulla neuron at a given time delay. In this model, the five small-field medulla neurons that are activated by the locally visual input, at different times of scanning, send their activities to a wide-field lobula neuron with a synaptic delay such that the lobula neuron receives all medulla input signals at the same instance (*i.e.* in the presented simulation the lobula neuron is maximally activated by the black vertical bar passing across the visual field from the left to right. Each underlying medulla neuron encodes the vertical bar in a different location of the visual field).

In more detail, the model sampled the image input into five sequential patches of 75x75 pixels with a speed of 0.1 m/s, equating to a lateral movement of 15 pixels between patches (an example of sequential scanning is shown in Figure 1B, see Method Section) (MaBouDi et al., 2021b). The green pixel intensities of each image patch modified the membrane potential of 75x75=5625 photoreceptors within the single eye of the simulated bee. These photoreceptor responses converged to 625 lamina neurons *via* recurrent neural connectivity. Lamina neurons provided post-synaptic connections to 250 small-field medulla neurons *via* simple feed-forward connectivity (Figure 1C, see Method Section). The medulla response from each image patch (of five patches that cover the image through the movement) is computed using a spiking neural model. Their spiking activities are then integrated into the synapses of their corresponding lobula neuron, with a time delay designed to ensure that the lobula neuron receives all the underlying medulla input signals at the same time. It is important to note that the proposed spatio-temporal coding is a simplification, as within the bee brain, this process would be achieved through dendritic and synaptic latency, or through intermediate neuron transmission within the medulla that is influenced by the non-associative learning in the visual lobe (Figure 1C, D). We hypothesised that connectivity in the medulla and lobula could be modified by exposing the visual lobe to a series of time-varying images while incorporating non-associative learning rules and efficient coding principle (see Method Section for details).

The neural representation of the visual inputs was subsequently transmitted and processed in the mushroom body, which serves as the visual learning centres of the bee brain (Ehmer and Gronenberg, 2002; Li et al., 2017; Paulk and Gronenberg, 2008) (Figure 1C). For simplicity, a single mushroom body output neuron (MBON) is exposed, where the firing rate of this neuron expresses the simulated bee's preference for any given visual input. By updating the synaptic weights within the mushroom body, we were able to train the neural network to associate visual patterns as either positive (resulting in low MBON firing rates) or negative (high MBON firing rates; see Discussion). After the visual network underwent acquisitive modification through non-associate learning and extensive exposure to natural images, the whole network, as a simulated bee, was trained and tested with a variety of pattern recognition tasks from the published literature (Benard et al., 2006; Dyer et al., 2005; Srinivasan, 2010, 1994; Zhang and Horridge, 1992), in particular with the "plus" and "multiplication sign" patterns used with the real bumblebees reported in MaBouDi et al., 2021 (Figure 2A). Finally, to assess the performance of the proposed active vision model in various pattern recognin tasks, we analysed the MBON as it is supposed to function as a pre-motor area. In this context, a lower response from the MBON to a specific pattern would indicate the bees' performance towards that particular pattern. Consequently, after multiple training trials, we presumed the MBON acts as a decision neuron, showing a lower response to the chosen pattern and higher response to the rejected pattern in the tests. Note that no reinforcement, or synaptic updates were implemented during the testing phase.

The model's performance in the same visual discrimination task is changed by scanning behaviour

To replicate the bee behavioural results seen in the published literature (Benard et al., 2006; Dyer et al., 2005; MaBouDi et al., 2021b; Srinivasan, 1994, 2010; Zhang and Horridge, 1992), we implemented the type of computational plasticity in the mushroom body circuitry that is necessary to mediate both appetitive and aversive value encoding. Here, we assumed that classical spike-timing-dependent plasticity (STDP) (Figure 7) modulated by dopamine governs the plasticity rule between the mushroom body Kenyon cells (KC) and extrinsic Mushroom Body Output Neurons (MBON) in the presence of negative (or unrewarded) patterns. Additionally, we introduced a novel plasticity rule using STDP (Figure S2B) modulated by octopamine, which we hypothesised leads to synaptic depression among KC-MBON connections. These different plasticity rules were employed to investigate the synaptic dynamics in response to positive and negative patterns (see Method and Discussion sections).

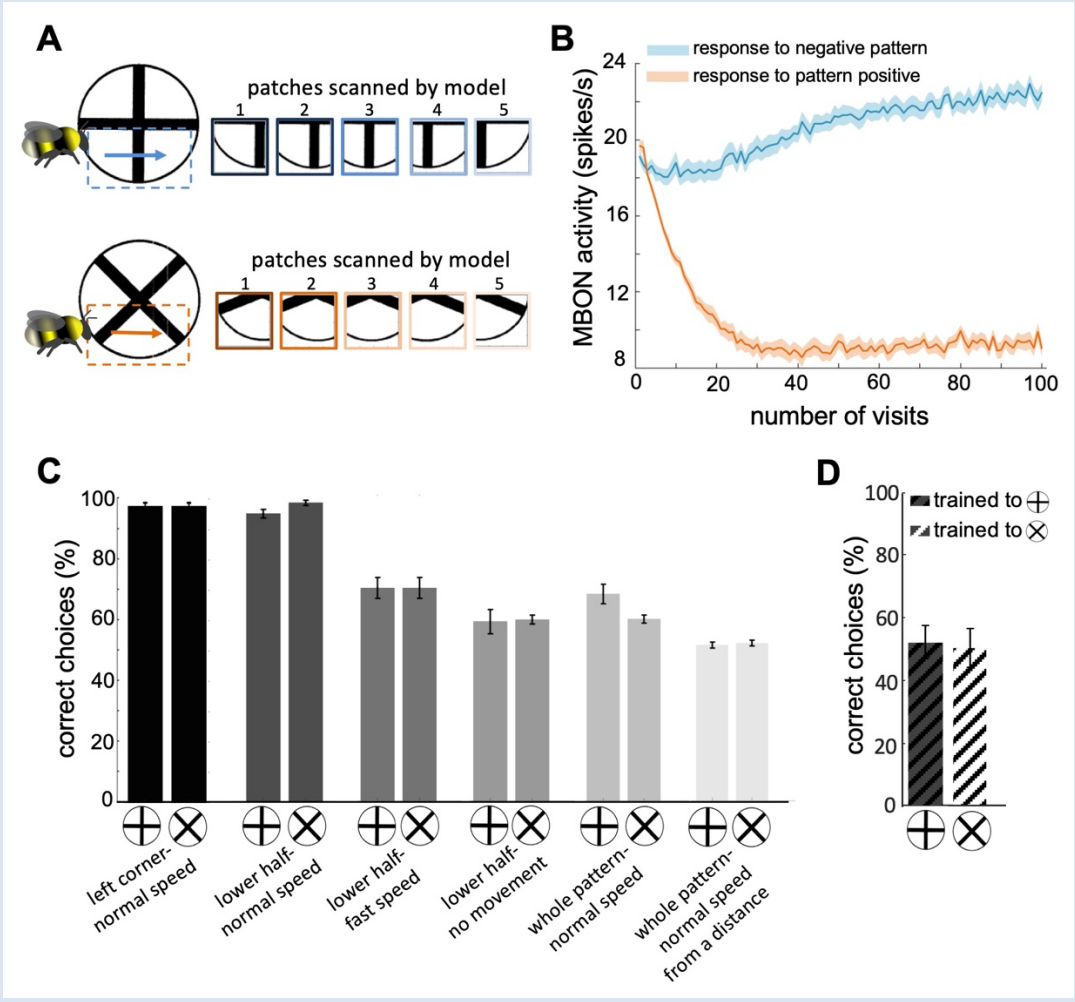


Figure 2. Simulated bees' performance in a pattern recognition task using different scanning strategies. Twenty simulated bees, with random initial neuronal connectivity in mushroom bodies (see Methods) and a fixed connectivity in the visual lobe that were shaped from the non-associative learning, were trained to discriminate plus from a multiplication symbol (100 random training exposures per pattern). The simulated bees scanned different regions of the patterns at different speeds. **(A)** Top and below panels show the five image patches sampled from the plus and multiplication symbols by simulated bees, respectively. It is assumed that the simulated bees scanned the lower half of the patterns with lateral movement from left to right with normal speed (0.1 m/s). **(B)** The plot shows the average responses of the MBON to rewarding multiplication and punishing plus pattern during training procedure (plus symbol rewarding, producing an Octopamine release by the reinforcement neuron, and the multiplication symbol inducing a Dopamine release). This shows how the response of the MBON to the rewarding plus was decreased while its response to the punishing multiplication pattern was increased during the training. The MBON equally responded to both multiplication and plus before the training (at number of visits =0). **(C)** The performance of the simulated bees in discriminating the right-angled plus and a 45° rotated version of the same cross (i.e. multiplication symbol) (MaBouDi et al., 2021b; Srinivasan, 1994), when the stimulated bees scanned different regions of the pattern (left corner, lower half, whole pattern) at different speeds: no speed 0.0m/s (i.e. all medulla to lobula temporal slices observed the same visual input), normal speed at 0.1 m/s and fast speed at 0.3m/s), and from a simulated distance of 2cm from stimuli (default) and 10cm (distal view). The optimal model parameters were for the stimulated bees at the default distance when only a local region of the pattern (bottom half or lower left quadrant) was scanned at a normal speed. **(D)** Mean performance (±SEM) of two groups of simulated bees in

discriminating the plus from multiplication patterns when their inhibitory connectivity between lobula neurons were not modified by non-associative learning rules.

Video 1: An example of bee's flight in recognising a plus sign pattern. The bee underwent training to receive 10 µl 50% sucrose solution (w/w) from the feeding tubes at the centre of the plus pattern. After carefully inspecting the lower half of the plus sign, the bee accurately chose the correct pattern (MaBouDi et al., 2021b).

Video 2: An example of bee's flight in recognising a multiplication sign pattern. The bee underwent training to receive 10 µl 50% sucrose solution (w/w) from the feeding tubes at the centre of the multiplication sign pattern. After carefully inspecting the lower half of the pattern, the bee accurately chose the correct pattern (MaBouDi et al., 2021b).

We first trained the model using a differential conditioning task in which the correct stimulus S+ is paired with reward while incorrect stimulus S- is delivered with punishment. Hereafter for simplicity, we use a subscript S to the label of valence of the pattern with + for positive and - for negative valence. In our simulations of associative learning, we update only those synaptic weights between KCs and MBON that correspond to the presented patterns (see Method section). To represent the learning task for different bees, we replicated the simulations, using different initial parameters (random neural connectivity between lobula and Kenyon cells and KCs-MBON connections). We then tested the performance of our model in a set of different visual task paradigms detailed below (Figures 2, 3). During the initial experiment where bees had to distinguish a plus from a multiplication sign, trained on the lower half of the plus (Figure 2A), the firing rate of the MBON decreases after presenting the plus (S+) and tends to increase after presenting S- (lower half of multiplication sign); whereas MBON responded equally to both plus and cross before training (Figure 2B). This indicates that the model can discriminate visual patterns S+ and S- using temporal coding and active scanning of patterns (Figure 2C). Conversely, the model with fixed random inhibitory connectivity in the lobula could not discriminate between the plus and the multiplication sign patterns (Figure 2D). This emphasizes the importance of the structured connectivity that emerges in the bee visual lobes by the non-associative learning (i.e. spatio-temporal receptive fields of lobula neurons), and subsequent performance of the simulated bees in visual learning tasks. In other words, rewarding patterns induce decreased extrinsic neuron responses, but result in higher responses for punished patterns (Figure 2B). This behaviour aligns with the biological observations in the alpha lobe of mushroom body PE1 neurons, which indicate a decrease in the response of MBON in the presence of the positive stimuli (Okada et al., 2007).

The replication of the plus and multiplication experiment with our simulated bees initially resulted in a poor performance compared to the real bees (Figure 2C), with the correct preference when trained on plus of 53%

(average result of 20 simulated bees) and just 52% for the reciprocal cross protocol. These simulations used the model parameters obtained from the bumblebee experiments: average flight speed whilst scanning of 0.1m/s, and a distance from the stimuli of 20mm. However, when the experiment was reconfigured for the simulated bees to scan only the bottom half of the patterns, or just the lower left corner (as seen with the real bees (MaBouDi et al., 2021), the correct choice performance increased to ≥96% and ≥98% respectively (Figure 2C). Increasing the scanning speed (*i.e.* increasing the separation between image patches) reduced the average performance to 70%. Similarly, simulated stationary bees only achieved 60% correct choices.

Neural network model of active vision bee behaviours in a variety of visual experiments

In this study, we evaluated our model (utilising scans from the lower half only) by comparing it with other bees' experiments reported in the literature. It is important to note that bees may exhibit variations in scanning behaviour under different patterns and training conditions (see Discussion section). We found our simulated bees could discriminate angled bars (Hateren et al., 1990), a 22.5° angled cross from a 90° rotated version (Srinivasan, 1994) and spiral patterns (Zhang and Horridge, 1992) (Figure 3A). If trained on five perpendicular grating patterns, the simulated bees correctly identified the correct novel grating and a single bar pattern variant (Figure 3B). Figure 3C shows that not only could the proposed neural network learn to identify the correctly oriented bar pattern, but also identify the rewarding pattern from a novel one (two circles). More importantly the model showed a lower preference (22%) for the negatively trained pattern to that of the same novel pattern. This validates the implementation of the rejection behaviour in the model, demonstrating that the model simultaneously learns the rewarding and aversive stimuli. This was further explored by training the neural networks with patterns that contained two oriented bars in each lower quadrant of the patterns (Figure 3D) (Benard et al., 2006; Stach et al., 2004; Zhang and Horridge, 1992). The simulated bees again discriminated the training patterns without difficulty (over 99% accuracy). However, they performed worse on a simplified variant of the patterns with an average result of just 61%. When tested with the original positive pattern and novel patterns that contained just one correct orientation the bees had a high preference for the correct stimulus. Equally, the simulated bees showed a preference, if not as dominant, for a pattern that contained just one correct feature versus the trained negative pattern; showing that the model was able to extract more than a single feature during its scan of the pattern. To provide a substantially more complex pattern recognition task, we replicated the facial recognition experiment performed on honeybees (Dyer et al., 2005). We trained the neural network with images of two human faces (Figure 3E). Similar to the performance seen in honeybees, our simulated bees were able to identify the positive trained face from the negative one, two novel faces and a caricature of a face. Both the real bees and our simulated ones were unable to discriminate the faces when rotated through 180°. This result shows that complex visual features can be condensed through spatiotemporal encoding in the lobula neurons into specific and distinct neuronal

representations needed for learning in a miniature brain of the bee. This result is quite remarkable but shows that complex visual features can be condensed through spatiotemporal encoding in the lobula neurons into specific and distinct neuronal representations needed for learning in a miniature brain of the bee.

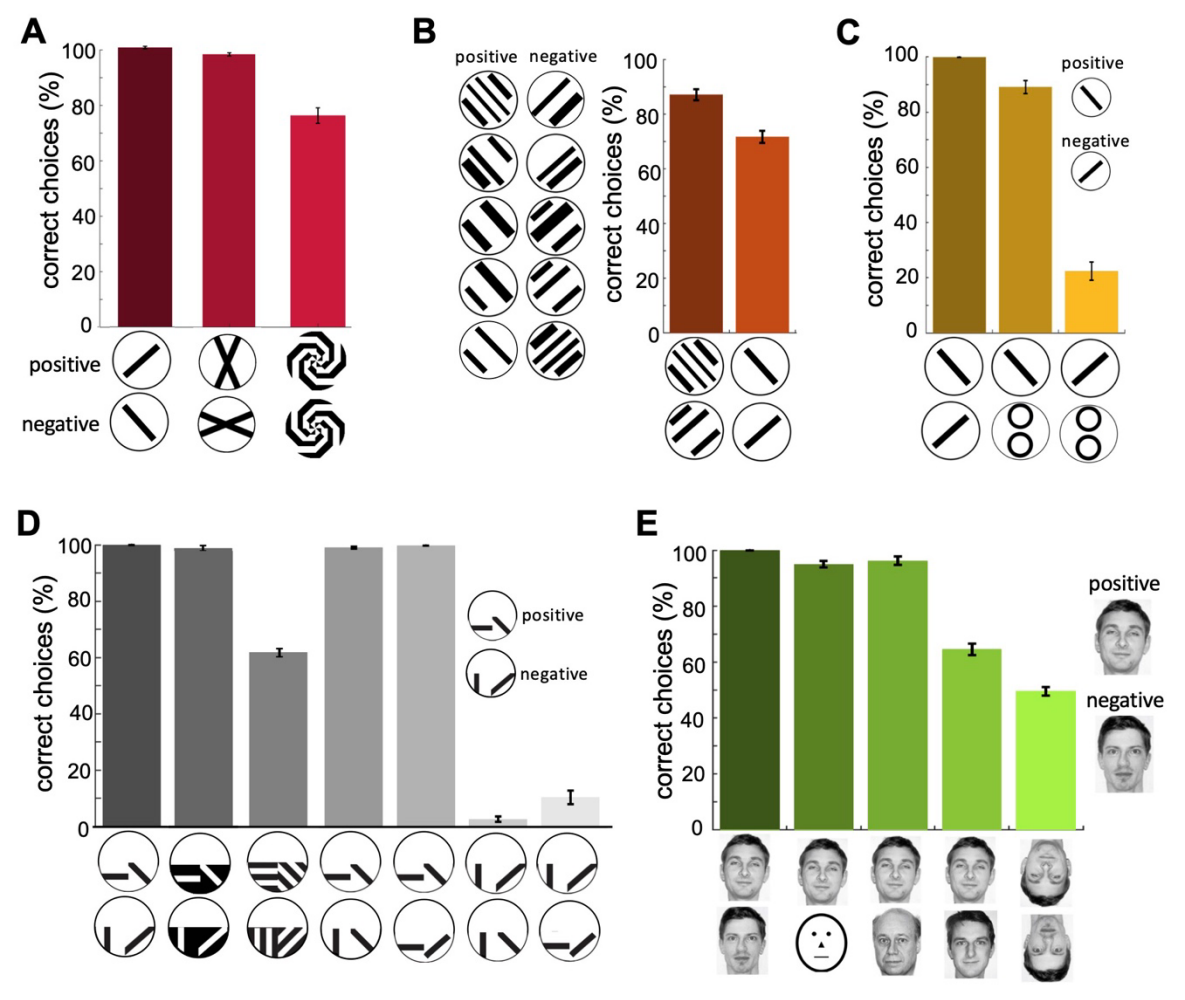


Figure 3. Minimal neural network performance to published bee pattern experiments. Twenty simulated bees, with random initial neuronal connectivity in mushroom bodies (see Methods), were trained to discriminate a positive target pattern from a negative distractor pattern (50 training exposures per pattern). The simulated bees' performances were examined via unrewarded tests, where synaptic weights were not updated (average of 20 simulated pattern pair tests per bee). Except for (A) all simulations were conducted at the default distance (2cm) and normal speed (0.1m/s) scanning the lower half of the pattern. **(A)** Mean percentage of correct choices (±SEM) in discriminating bars oriented at 90° to each other, 25.5° angled cross with a 45° rotated version of the same cross, and a pair of mirrored spiral patterns (MaBouDi et al., 2021b; Srinivasan, 1994). The simulated bees achieved greater than chance performances. **(B)** Performance of simulated bees trained with a generalisation protocol (Benard et al., 2006). Trained to 6 pairs of perpendicular oriented gratings (10 exposures per grating). Simulated bees then tested with a novel gating pair, and a single oriented bar pair. The simulated bees performed well in distinguishing between the novel pair of gratings; less well, but still significantly above chance, to the single bars. This indicates that the model can generalise the orientation of the training patterns to distinguish the novel patterns. **(C)** Mean performance (±SEM) of the simulated bees in discriminating the positive orientation from negative orientation. Additionally, the performance in recognising the positive

orientation from the novel pattern, and preference for the negative pattern from a novel pattern. Simulated bees learnt to prefer positive patterns, but also reject negative patterns, in this case preferring novel stimuli. **(D)** Performance of simulated bees trained to a horizontal and -45° bar in the lower pattern half versus a vertical and +45° bar (Stach et al., 2004). The simulated bees could easily discriminate between the trained bars, and a colour inverted version of the patterns. They performed less well when the bars were replaced with similarly oriented gratings, but still significantly above chance. When tested on the positive pattern vs a novel pattern with one correctly and one incorrectly oriented bar the simulated bees chose the positive patterns (fourth and fifth bars), whereas with the negative pattern versus this same novel pattern the simulated bees rejected the negative pattern in preference for the novel pattern with single positive oriented bar (two last bars). **(E)** The graph shows the mean percentage of correct choices for the 20 simulated bees during a facial recognition task (Dyer et al., 2005). Simulated bees were trained to the positive (rewarded) face image versus a negative (non-rewarded) distractor face. The model is able to recognise the target face from distractors after training, and also to recognise the positive face from novel faces even if the novel face is similar to the target face (fourth bar). However, it failed to discriminate between the positive and negative faces rotated by 180°.

Neural network response results during pattern learning

The lobula neurons were additionally configured to laterally inhibit each other, allowing for non-associative learning (synaptic plasticity in the absence of reward) by employing synaptically local rules between lamina, medulla and lobula neurons (see Method section). The feed-forward weights from lamina to medulla neurons were updated with Oja's implementation of Hebb's rule (Oja, 1982) while symmetric inhibitory spike-timing-dependent plasticity (iSTDP) rule was implemented in synaptic connectivity of the lateral inhibitory lobula neurons (Vogels et al., 2011). These rules drive the network to efficiently represent the visual input with a limited number of lobula neuron activations (Figure 4B) (see Discussion). Figure 4A illustrates the obtained receptive fields of lobula neurons representing the spatiotemporal orientation-selectivity obtained after training with 100 flower and natural images (50,000 time-varying image patches). Each square represents one of the 50 lobula neurons, with the heat map representing the synaptic weights of the associated lamina neurons (connected via the medulla neurons). As representing spatiotemporal results in Figure 4A is inherently unintuitive, an example of two lobula neurons is shown below the matrix with the lamina synaptic weights for each of the five medulla neurons (see temporal dynamic of obtained RFs in Video 3 for all 50 lobula neurons, and Figure 4 - figure supplement 1 and Video 4 for all 100 lobula neurons). The receptive fields of a lobula neuron contain an elongated 'on' area (positive synaptic weights), next to an antagonistic "off" area (negative synaptic weights) that are mostly arranged in parallel in a specific orientation. The proportion of both 'on' and 'off' areas continuously change from one time-delayed instance of medulla responses to the next. In the first example, the lobula cell with such a receptive field responds best to a 135° angled bar moving in the direction orthogonal to the orientation of the 'on' or 'off' areas and produces little or no responses to other orientations (Figure 4*C*). Our population of 50 lobula neurons show specificity to orientation and direction, similar to the neuronal responses observed in the brain of bees and other insects (James and Osorio, 1996; Paulk et al., 2008; Seelig and Jayaraman, 2013; Yang and Maddess, 1997).

280

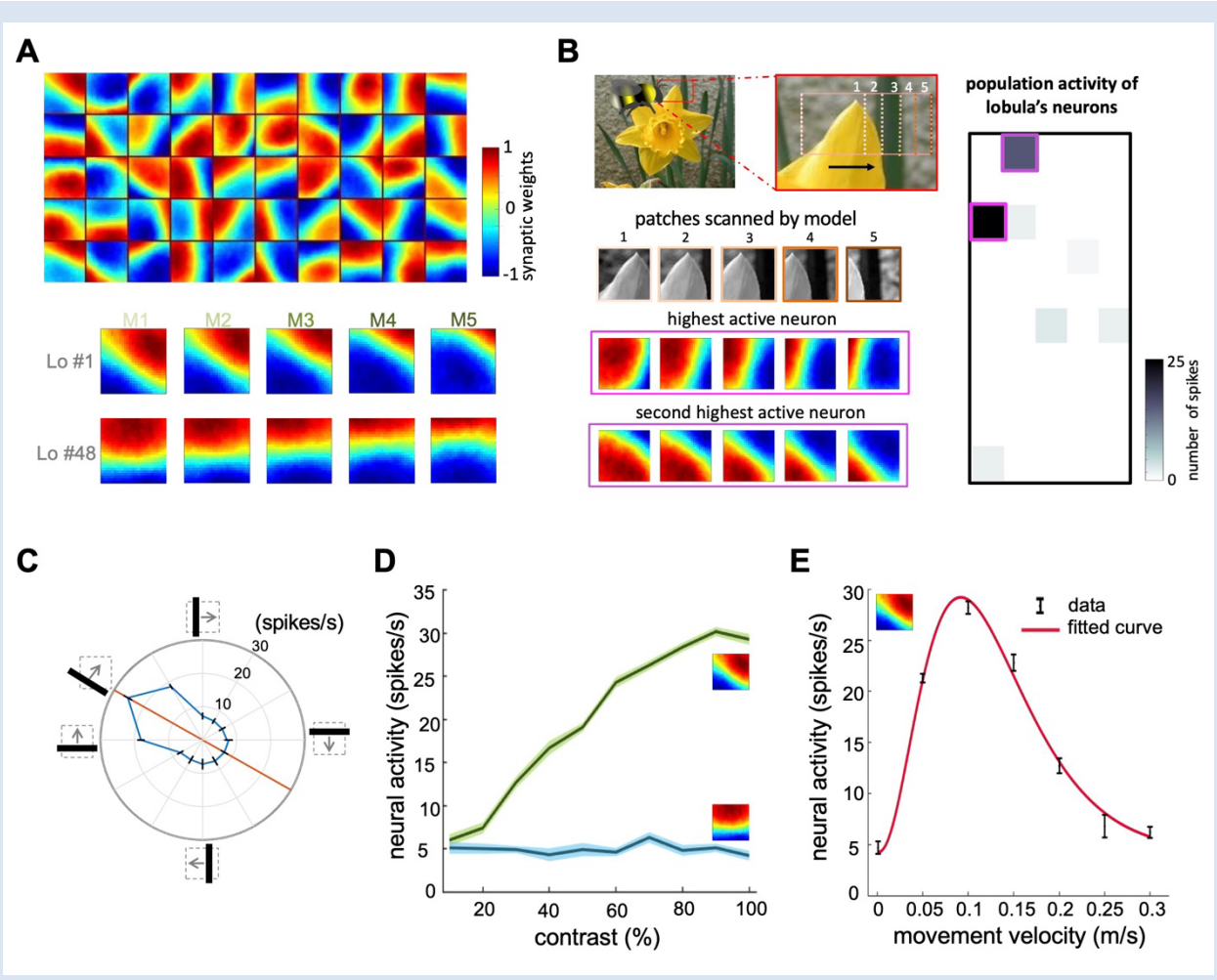


Figure 4. Neural responses of the simulated bee model to visual patterns. (A) Top: each square in the matrix corresponds to a single time slice of the obtained spatiotemporal receptive field of a lobula neuron (5x10 lobula neurons) that emerged from non-associative learning in the visual lobes after exposing the model to images of flowers and nature scenes (see Video 3). Bottom: spatiotemporal receptive field of two example lobula neurons are visualised in the five-time delay slices of the matrices of synaptic connectivity between lamina and five medulla neurons (See Figure 1D). The lobula neuron integrates signals from these medulla neurons at each of five time periods as the simulated bees scan a pattern (time goes from left to right). Blue and red cells show inhibitory and excitatory synaptic connectivity, respectively. The first example lobula neuron (#1) encodes the 150° angled bar moving from lower left to the upper right of the visual field. The second example lobula neuron (#48) encodes the movement of the horizontal bar moving up in the visual field. **(B)** An example of an image sequence projected to the simulated bee' eye with lateral movement from left to right. Below shows the five images patched sampled by the simulated bee. The right side presents the firing rate of all lobula neurons responding to the image sequence. The spatiotemporal receptive field of two highest active neurons to the image sequence are highlighted in purple. **(C)** The polar plot shows the average orientation selectivity of one example lobula neuron (#1) to differently angled bars moving across the visual field in a direction orthogonal to their axis (average of 50 simulations). This neuron is most sensitive to movement when the bar orientation is at 150°. **(D)** The spiking response of the lobula neuron to the preferred orientation raised as the contrast was increased, whereas the response of the lobula neuron to a non-preferred orientation is maintained irrespective of contrast. **(E)** The average velocity-sensitivity curve (±SEM) of the orientation-sensitive lobula neuron (#1) is obtained from the responses of the lobula neuron to optimal

(angle of maximum sensitivity) moving stimuli presented to the model at different velocities. The red line shows the Gamma function fitted to the data.

Video 3: Temporal dynamic of receptive field of 50 lobula neurons obtained from non-associate learning procedure and active scanning.

For clarity, Figure 4B shows how a sequence of image patches simulating a horizontal scan is processed, and results in a limited number of lobula neuron responses. This means that the activity of lobula neurons is decorrelated and relatively sparse as a result of non-associative learning mechanisms in the visual lobe (See Discussion). Interestingly, the receptive field of two first active lobula neurons captured the visual structure of the flower's petal that was scanned by the simulated bee (*i.e*. one is matched with left 45 angled edge of the flower's petal and the other is almost matched with the right angled edge of the flower's petal). This shows that our model can extract the different visual features of the input with a minimum number of filters (lobula neurons). To further demonstrate the characteristics of the 150° sensitive lobula neurons (as one example), the spiking activity of the neuron is calculated in the response to a set of oriented bars as they were moved in the direction orthogonal to their orientation. As expected, the neuron is maximally sensitive to a moving 150° bar (26 spikes/sec), it still responds to a horizontal bar and a moving 120° bar (18Hz) but has limited responses (above base firing rate) to the other stimuli (Figure 4C). Consistent with neural observations (Yang and Maddess, 1997), the firing rate response of our lobula neurons, at their 'preferred' orientations, rises as the contrast is increased. Conversely, their responses to the non-preferred orientations are constant irrespective of contrast (Figure 4D). Interestingly, these lobula neurons are also velocity-sensitive (Figure 4E). Thus, each lobula neuron optimally responds to a specific orientation and velocity. This demonstrates that our model captures many quantitative features of lobula neurons (Paulk et al., 2008; Yang and Maddess, 1997).

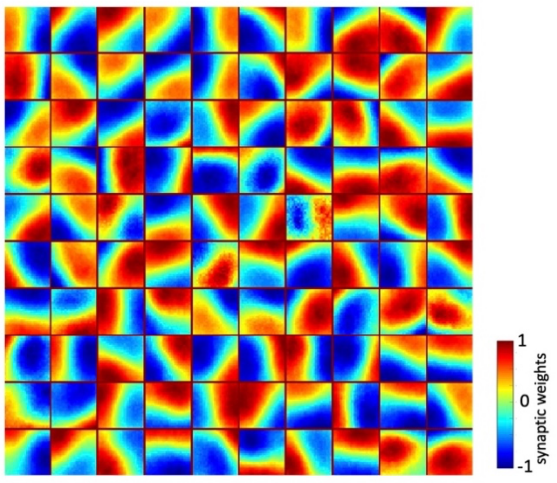


Figure 4 - figure supplement 1: The spatiotemporal receptive field of lobula neurons emerged from the non-associative learning if the number of lobula neurons is set to 100 (see Video 4).

Video 4: Temporal dynamic of receptive field of 100 lobula neurons obtained from non-associate learning procedure and active scanning.

What is the minimally sufficient circuitry required for active vision in bees?

As reported above, the model is able to accomplish a variety of pattern recognition tasks (Figures 2, 3). We asked whether our neural networks could perform with a very limited number of lobula neurons that transfer visual information to the mushroom body. Hence, we conducted the non-associative learning process with varying number of lobula neurons, setting them to either 4, 16, or 36 (the original model had 50 lobula neurons). Then the visual network was trained using the same set of natural images and protocol defined for the original model. Interestingly, the non-associative learning process led to the emergence of distinct spatio-temporal structures in the lobula neurons. We found that the variability between the spatiotemporal receptive field of lobula neurons is reduced by decreasing the number of lobula neurons (Figure 5A and Videos 5, 6 ,7). The model cannot encode the spatio-temporal structure inherent in the training patterns if the model is limited to 4 neurons, only vertical and horizontal receptive fields were created (Video 5). As expected, the performance of the model decreased by reducing the number of lobula neurons. However, the model with 16 lobula neurons can still solve the discrimination between plus and multiplication signs, and solve a difficult visual task above chance (Figure 5B, C). This finding highlights how the spatio-temporal encoding in the visual lobe enhances the model's capacity to represent the visual environment using fewer neurons, compared to what is typically necessary in a minimal model.

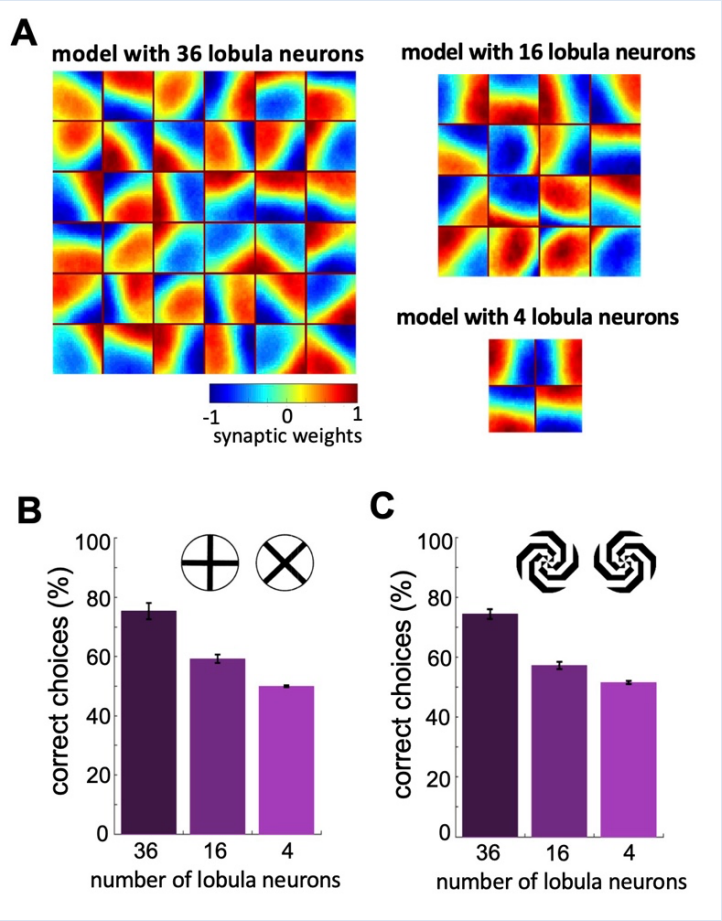


Figure 5. Minimum number of lobula neurons that are necessary for pattern recognition. (A) Obtained spatiotemporal receptive field of lobula neurons when the number of lobula neurons were set at 36, 16 or 4 during the non-associative learning in the visual lobe (See Figure 5A). This shows the models with lower number of lobula neurons encode less variability of orientations and temporal coding of the visual inputs (see Videos 5,6) **(B & C)** The average correct choices of the three models with 36, 16 or 4 lobula neurons after training to a pair of plus and multiplication patterns (B) and mirrored spiral patterns (C). The model with 16 lobula neurons still can solve pattern recognition tasks at a level above chance. It indicates that only 16 lobula neurons that provide all inputs to mushroom bodies are sufficient for the simulated bees to be able to discriminate between patterns.

Video 5: Temporal dynamic of receptive field of 4 lobula neurons obtained from non-associate learning procedure and active scanning (compare it to Video 3).

Video 6: Temporal dynamic of receptive field of 16 lobula neurons obtained from non-associate learning procedure and active scanning (compare it to Video 3).

Video 7: Temporal dynamic of receptive field of 36 lobula neurons obtained from non-associate learning procedure and active scanning (compare it to Video 3).

Moreover, to explore the effect of inhibitory neurons within the visual lobe on the output of lobula neurons, the model was trained using the same protocol, but the synaptic weights of the inhibitory connections to the lobula neurons were not updated during exposure to the training images. These fixed inhibitory connections

caused a limitation in shaping the population of lobula neurons to encode moving orientations (Figure 6 and Video 8). It indicates that the presence of inhibitory interneurons in the visual lobe plays a crucial role in facilitating an efficient representation of the visual environment.

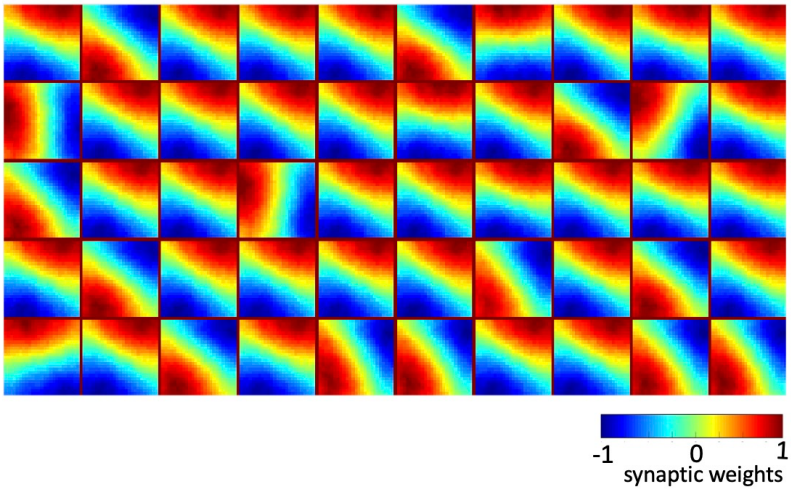


Figure 6. The role of lateral inhibitory connections between lobula neurons. Obtained spatiotemporal receptive field of lobula neurons when the lateral inhibitory connectivity between lobula neurons is fixed during the non-associative learning (see Video 8 and Figure 4A & Video 3).

Video 8: Temporal dynamic of receptive field of 50 lobula neurons when the lateral inhibitory connectivity between lobula neurons is fixed during the non-associative learning (compare it with Video 3)

Taken together, these findings demonstrate that our assumption regarding non-associative plasticity in the visual lobe can successfully replicate the neural responses of lobula neurons across various patterns and conditions. This results in a sparse and uncorrelated representation of the visual input, which is advantageous for subsequent learning processes in the mushroom body. Importantly, these results align closely with theoretical studies (see Discussion section), further supporting the effectiveness of the active vision in capturing the underlying principles of information encoding in the insect visual system.

Discussion

In this study, we aimed to gain insights into the computational requirements for visual pattern recognition by investigating a minimal neural network. To achieve this, we leveraged the flight behaviours of bees during their active scanning of visual patterns and developed a novel model inspired by the insect's visual system. Through simulations, we examined how the visual environment is represented through the spatio-temporal responses of a small population of neurons in the lobula region in the visual lobe. By incorporating non-

associative learning into our model, we discovered that it can effectively shape the connectivity within the visual lobe and generate an efficient representation of the input. This self-organization process led to the emergence of orientation-selective cells in the lobula, which played a vital role in encoding the complex visual environment. Our model of the bees' visual system, and subsequent simulations, demonstrate that the complex visual environment can be condensed into spatiotemporal representations, expressed through the firing rate responses of a small population of lobula neurons sensitive to specific orientations and velocities. Not only are these limited representations capable of discriminating the plus and multiplication patterns used in our behavioural experiments (MaBouDi et al., 2021b), but can also generalise from trained patterns to novel stimuli, and even provide accurate results in a task of human face recognition, indicating its potential for broader applicability. Our findings also suggest that the movement of bees, or their active vision, may play a crucial role in their ability to efficiently analyse and encode their environment. The spatio-temporal encoding within the visual lobe appears to be a key mechanism employed by bees to achieve this efficient information representation. Overall, our study sheds light on the necessary computational requirements for visual pattern recognition, highlighting the significance of active vision and spatio-temporal encoding within the insect's visual system. These insights have implications not only for understanding the information processing capabilities of bees but also for inspiring the development of novel computational models for visual recognition tasks.

The question of how animals deal with noisy and complex natural world has a long history (Barlow, 1961; Gibson, 1979; Menzel and Giurfa, 2006; Srinivasan, 2010). The efficient coding hypothesis is one of the important theoretical hypotheses in this field that refers to the phenomenon that the early visual system compresses inputs into a more efficient form to transmit relevant visual information with maximum information to higher order brain regions (Barlow, 1961). According to this hypothesis, the responses of a single cell in the visual system to the natural environment should entirely utilise its output capacity (e.g., maximum firing rate) and the population responses of different neurons to the natural signals should be statistically independent (Simoncelli and Olshausen, 2001). Bees present a wide range of complex behaviours over small scales of inspections and large scales of navigation (Menzel, 2012; Srinivasan, 2010) despite their limited computational resources. Thus, bees are an appropriate model to explore the effect of ecological constraints in the neural computation underlying cognition. The non-associative model presented in this study supports the efficient coding hypothesis (Figures 4 & 5). The suggested model works as a linear generative model that successfully replicates the receptive fields of cells in the lobula (Barlow, 1961; Olshausen, 2003). After training, the correlation among activity of lobula neurons to spatiotemporal naturalistic signals is highly reduced. Further, only a limited number of lobula neurons respond to specific visual stimuli (Figure 4). This indicates that the model removes redundancy in the natural scenes such that the receptive fields of lobula

cells often encode perceptually salient features in the natural scene and transfer the incoming signals into a more efficient form.

In this study, we use the type of natural scenes with a statistical structure similar to those that visual systems have adapted to over evolutionary time periods (Geisler, 2008; Hyvärinen et al., 2009; Simoncelli and Olshausen, 2001). Also, since the main goal of bee navigation and foraging is finding food from a large variety of potential flower resources, our non-associative network was trained with a set of different flower images. As with all theoretical models this is a simplification, as real bees traverse a 3D environment viewed through a 270° field of view. Here, we assume that formation of receptive fields would be equivalent to that of our 2D simulations. However, further studies are necessary to refine and expand our model based on a more comprehensive understanding of the function and structure of the bee eye components (Juusola et al., 2017; Taylor et al., 2019). Moreover, investigating the neural mechanisms underlying visual learning in the bee brain will allow us to fine-tune our model's architecture and parameters, leading to a more faithful representation of the bee visual system. Through continued research and collaboration, we plan to refine our model and gain deeper insights into the remarkable capabilities of the bee's visual perception.

Our findings align with previous studies on bumblebees' discrimination of plus and multiple sign patterns (MaBouDi et al., 2021b), indicating improved performance of our model when focusing on scanning the lower half of patterns at specific velocities. However, bees exhibit variations in scanning behaviour under different patterns and training conditions (Giurfa et al., 1999; Guiraud et al., 2018). It has been shown that honeybees and bumblebees can solve visual tasks by extracting localized or elemental features within patterns, adapting their discrimination strategies accordingly (Giurfa et al., 1999; MaBouDi et al., 2021b; Stach et al., 2004; Stach and Giurfa, 2005). Recent analysis of bee flight paths supports these findings, revealing how honeybees can successfully solve tasks by selectively scanning specific elements in stimuli (Guiraud et al., 2018). This suggests that bees may develop tailored flight manoeuvres during training, optimizing their scanning behaviour to extract maximum visual information based on patterns and protocols. Although our model highlights the importance of studying active vision, it is essential to recognize that further investigations are needed to explore the optimal flight scanning behaviour in bees. Additionally, incorporating an adaptive vision-motor loop into our models will enable a more comprehensive understanding of active vision in insects. These advancements will provide valuable insights into how insects perceive and interact with their visual environment, ultimately enhancing our understanding of the mechanisms underlying active vision.

The results of our model suggest that passive visual exposure to natural images modifies the connectivity in the visual lobes and leads to better ability in pattern recognition (Figures 2, 3). These developed synaptic

connections emerged independent of the initial connectivity that each simulated bee had upon creation. We assume that the specific visual experiences that real bees are exposed to in early life define the individual capabilities of each bee's visual representation, which may influence their later performance in behavioural tasks (Hertel, 1983, 1982; MaBouDi et al., 2017; Vetter and Visscher, n.d.). There is direct empirical evidence for such neural developmental processes in olfactory systems of bees, where early passive exposure improved the bees' later ability in odour discrimination (Arenas and Farina, 2008; Locatelli et al., 2013). Our previous research of olfactory coding demonstrated that the iSTDP learning rule can create specific connectivity in the sensory system and increase the separability of odour representations in the antennal lobe outputs (MaBouDi et al., 2017). A similar mechanism was developed here between lobula neurons, such that only a limited population of lobula neurons are excited for particular visual inputs, providing sparse and distinct outputs to the mushroom body learning centres (Figure 3B). The obtained receptive fields of lobula cells with the fixed lateral connectivity shows the inhibition is required for orientation selectivity and temporal coding in the visual lobe (Fisher et al., 2015). Our results highlight the important functions of inhibitory connections within the visual lobes. Accordingly, our model predicts that bees with less experiences of visual processing at the early stage of life are worse in the learning and memory of visual tasks compared with bees with rich visual experiences. However, further behavioural and neurobiological studies are required to assess this prediction.

Mushroom bodies are critical centres for associative learning and memory in insects (Heisenberg, 2003; Menzel, 2012). Synapses between Kenyon Cells and extrinsic mushroom body neurons obey a Hebbian STDP rule (Cassenaer and Laurent, 2007; Markram et al., 1997). However the STDP rule alone cannot maintain associate learning (Abbott and Nelson, 2000; Meeks and Holy, 2008). Associative learning in insects appears to rely on the neurotransmitters octopamine and dopamine, to reflect unconditioned signals for appetitive and aversive valances (Cognigni et al., 2018; Hammer, 1993; Hammer and Menzel, 1995; Perry and Barron, 2013; Schwaerzel et al., 2003). These are released into the mushroom body lobes where Kenyon cells connect to MB output neurons (MBON) (Burke et al., 2012; Okada et al., 2007; Strube-Bloss et al., 2011). Cassenaer and Laurent, using in-vivo electrophysiology in locusts, reported the depressive action of octopamine on synapses underlying STDP rule that leads to a lower response for MOBNs in the presence of octopamine (Cassenaer and Laurent, 2012). Hence, following this observation, we model associative learning of pairing the positive pattern with the reward by the octopamine modulation of STDP (Equation 4; Figure 7). Here, both temporal ordering of pre- or post- synaptic spikes depress the synaptic connection between Kenyon cells and the MBON.s Also, the synapses are updated when the negative patterns are paired with the punishment following the classical STDP (Equation 3; Figure 7). This combination results in a complex interplay between synaptic changes and reinforcer signals and enriches the model to not only learn to correctly choose the positive patterns, but also learn to reject incorrect patterns (Figure 2B, 3C). The changes to the synaptic weight

of connections results in a decreasing response of MBON to the positive patterns during the associative learning which is consistent with the PE1 extrinsic neuron in the honeybee brain that exhibits a lower response to the positive patterns (Okada et al., 2007) (Figure 2B). However, further studies are required to investigate the novel combination of octopamine and dopamine modulation of STDP that is introduced in this study. Together the non-associative learning in the optic lobes and supervised learning in the mushroom bodies produced a model capable of not only discriminating simple patterns but also generalisation (Figure 3B), and correct judgments in conflicting stimulus experiments (Figure 3D). However, the real power of this approach is exemplified in the facial recognition task (Figure 3E). Here, the complexity of the human face is reduced to a number of sparse lobula neuron activations that can be learnt by the mushroom bodies. But more interesting is that the spatiotemporal receptive fields formed during non-associative learning respond differently for different faces, allowing fine differences to still be sufficiently encoded. Real bees rarely have to discriminate between human faces, but these same processes undoubtedly aid bees in selecting rewarding flowers without requiring a complex visual memory within their miniature brains.

Recent studies showed that bees can sometimes use a more efficient, less-cognitively demanding strategy to solve a cognitive task (Cope et al., 2018; Guiraud et al., 2018; Langridge et al., 2021; MaBouDi et al., 2023, 2021a, 2020b; Roper et al., 2017; Vasas et al., 2019). Most of bees' responses to different cognitive tasks can be described by a simple neural network. For instance, Roper et al. (2017) suggest that reliable generalisation of visual information can be achieved through simple neuronal circuitry that is biologically plausible and can be accommodated in a small bee brain (Roper et al., 2017). Also, a mathematical model of colour processing in the bee brain propose that the diversity of colour-sensitive responses can be explained using a simple model by the assumption that these neurons receive randomly weighted inputs from all receptor types, and that this type of neural organisation is likely implemented during neural development and experience-dependent manner (MaBouDi et al., 2020b; Vasas et al., 2019). Together, the behavioural and computational experiments developed during our research emphasise the fundamental roles of exploring the mechanisms of cognitive abilities in animals with miniature brains and designing minimum neural networks to understand the requirements of certain cognitive tasks. A mechanistic investigation of how bees parse natural environments provides basic principles for current challenging problems in designing autonomous robots. Indeed, these computational shortcuts that have evolved for billions of years will enable us to develop more efficient artificial intelligence, capable of solving specific problems much more effectively than humans and current artificial intelligence (de Croon et al., 2022; Webb, 2020).

Materials and Methods

Network topology of active vision model

The model architecture of the bee visual pathway is shown in Figure 1. A bumblebee has a pair of compound eyes that are composed of ~5,500 ommatidia (Spaethe and Chittka, 2003; Streinzer et al., 2013). Each eye contains three different types of photoreceptors, short, medium and long wavelength sensitive peaking in the UV, blue and the green respectively (Menzel and Blakers, 1976; Skorupski et al., 2007). Since the green photoreceptors are those that predominantly mediate visual pattern recognition (Giger and Srinivasan, 1996; Spaethe et al., 2001), we modelled that 75x75 green photoreceptors in one eye component are activated by the pixel values of the input pattern. Photoreceptors then project to 625 (25x25) neurons in the lamina, which is the first centre of visual processing. In this model, one lamina neuron, r_l^{La}, receives and sums inputs from nine neighbouring photoreceptors placed in neighbouring ommatidia, as $r_l^{La} = f(\sum_{p=1}^{P} r_p\,;\, a\,, b)$. Here, $f(x; a, b) = A_0\,/(1 + exp(mx + b)$ is the activation function where $A_0 = 1$ is the maximum activity of the lamina neurons and parameters $m = 1$ and $b = 0$ control the shape of the activation function f.

Each medulla neuron is activated by the summation of activity of the lamina neuron via the synaptic connectivity W. Each spiking neuron in this study follows the integrate-and-fire model. The dynamics of the subthreshold membrane potential of a neuron, $u(t)$ is described by the following standard conductance-based leaky integrate-and-fire model: $\tau \frac{du(t)}{dt} = -u(t) + R.I(t)$, where $R = 10$ and $\tau = 10ms$ are the resistance and membrane time constant of the neuron respectively. Here, the input $I(t)$ exhibit the total synaptic input to the cell from presynaptic neurons.

The membrane potential is reset to the base activity, $v_0 = -80\,mV$, if it exceeds the threshold, $V_T = 0\,mV$. However, the input of the m −the medulla neuron is calculated $I_m^{Me} = \sum_{l=1}^{L} W_{m,l}\; r_l^{La}$. The value $W_{l,m}$ specifies the strength of a synaptic input from the l −th lamina neuron to the m −the medulla neuron. To model the variability of neural responses, a signal noise generated by Poisson distribution was added to the output of the neuron.

We propose a temporal coding model between the medulla and lobula neurons. Each wide field lobula neuron receives synaptic input from M small field medulla neurons with delay T(Figure 1D) (M corresponds to the number of frames in the input of the model). These small field medulla neurons are activated from M different regions of the image patch via the pathway passing through the photoreceptors and lamina neurons (Figure 3B). While each medulla neuron is affected by 1/M of the pattern, a lobula neuron is activated by the whole pattern. M medulla neurons that are sampled from a selected region with delay T_0 show the neural representation of the bees' scanning behaviour in front of the pattern. The medulla neurons send their spiking responses to a wide-field lobula neuron with a synaptic delay such that all signals activate the lobula neuron in the same instance. The parameters of the model's scanning behaviour (viewing distance: 2cm, flight speed 0.1m/s) are obtained from bees whilst inspecting stimuli (MaBouDi et al., 2021b).

The lobula neurons are laterally interconnected by the inhibitory neurons via the synaptic connectivity $Q = [q_{i,j}]$, where $q_{i,j}$ represents the lateral connectivity between i −th and j −th lobula neurons. This connectivity along with $W_{m,l}$ are updated during a non-associative learning process (see next subsection). The neurons in the lobula region send excitatory signals to randomly selected Kenyon cells (KCs) in the mushroom bodies through the synaptic connectivity matrix (Caron et al., 2013; Szyszka et al., 2005) $S = [s_{o,k}]$. The positive values $s_{o,k}$ exhibits the strength synaptic weight from o-th lobula neuron to k-th KC. KCs in the network have sparse activity, meaning they are selective to particular image features (i.e. each pattern activates less than 5% of Kenyon cells(Honegger et al., 2011)). All Kenyon cells project to a single mushroom body Output Neuron (MBON), which is the final output of the model. The input of the of the MBON, I_{MBON}, is computed by the KC-MBON connections D such that $I_{MBON} = \sum_{k=1}^{K} D_k r_k^{KC}$, where r_k^{KC} is the spiking activity of the k-th KCs. Finally, a reinforcement neuron makes reinforcement-modulated connections with the KCs and MBON in the presence of the positive and negative patterns (see the next section).

Training the network via a non-associative learning

We trained the model on 50,000 time-varying patches randomly selected from 100 flowers and natural scenes. At each step of training, a set of five patches with size 75x75 pixels, selected by shifting 15 pixels over the image from the left or right or the reverse orientation (Figs 1B, 2A), was considered as the input of the model; this would correspond to the process of bees' scanning part of the image. Using the network described above, the number of spikes from each lobula neuron are counted separately for each set of patches scanned through each movement. We start the training with all inhibitory connection strengths Q, where its elements are randomly generated from a uniform distribution between 0 and 1. The feed-forward synaptic weights are initialized with Gaussian white noise $\mathcal{N}(0, 1)$. The neural responses random time-varying patches evoked by the images were used to update the connections strengths Q and W, simultaneously (see Discussion section).

After the image presentation, the feed-forward weight W is updated according to Oja's implementation of the Hebbian learning rule (MaBouDi et al., 2017; Oja, 1982) via

$$\Delta W_{i,j} = \gamma r_j^{Me}(r_i^{La} - r_j^{Me} W_{i,j}) \qquad \text{(Equation 1)}$$

Here, the r_j^{Me} and r_i^{La} represent the activities of the j-th medulla and i-th lamina neurons, respectively. The positive constant γ defines the learning rate.

At the same time of processing, the lateral inhibitory connectivity in the lobula is modified by inhibitory spike-time-dependent plasticity (iSTDP) (Vogels et al., 2011). Here, we model *non-associative learning* in lobula by a

symmetric iSTDP between presynaptic of the inhibitory neurons and postsynaptic lobula neurons. In this learning rule, both temporal ordering of pre- or post- synaptic spikes potentiate the connectivity and the synaptic strength of j-th inhibitory neuron onto i-th lobula neuron ($Q_{i,j}$) is updated as follows:

$$\Delta Q_{i,j} = \eta(r_i^{Lo} * r_j^{In} - \alpha) \qquad \text{(Equation 2)}$$

where r_i^{Lo} and r_j^{In} exhibit the mean firing rate of the lobula and inhibitory neurons, respectively. The depression factor α controls the target activity rate of the lobula neurons. Here, η is the learning rate. To simplify, a one-to-one connection between the inhibitory and lobula neuron is assumed in the model such that the activity of the j-th inhibitory neuron is equal to the activity of the j-th lobula neuron. The training is terminated when the synaptic weights over time are changed less than a very small threshold (0.001).

Associative learning in Mushroom Bodies

To verify if the lobula neurons can reproduce empirical behavioural results in different visual tasks, the model is enriched with a supervised learning process in the mushroom bodies. When the training process of the non-associative learning is terminated, we use a reward-based synaptic wright modification rule in KCs-MBON connection (D), such that, if a stimulus is rewarding (*i.e.* positive), the corresponding synapses between activated neurons will be weakened while for a stimulus paired with punishment (*i.e.* negative), activated synapses are strengthened (Cassenaer and Laurent, 2012) (see Discussion section). The model behaves as the activity of mushroom body neurons in decreasing their firing rate in responding to the positive stimuli during training (Okada et al., 2007). In this model, a single reinforcement neuron modulated strengths of synaptic connectivity at the output of the KCs in response to both reward and punishment. In the presence of the negative patterns, the synaptic strengths from the KCs to the MBON are modified, and modulated by dopamine, based on the classical STDP (Song et al., 2000; Zhang et al., 1998) (Figure S3A):

$$STDP_{Dop}(\Delta t) = \begin{cases} A\, e^{-\Delta t/\tau}, & \Delta t > 0 \\ -A\, e^{\Delta t/\tau}, & \Delta t < 0 \end{cases}, \qquad \text{(Equation 3)}$$

where $\Delta t = t_{post} - t_{pre}$ implies the difference between the spike time of pre- and post- synaptic neurons. Further, applying the synaptic plasticity rule modulated by octopamine (octopamine modulated STDP) observed in the presence of rewarding stimuli to the synapses between KCs and MBON (Cassenaer and Laurent, 2012), the change in synaptic weight can be summarized as (Figure 7):

$$STDP_{OCT}(\Delta t) = \begin{cases} -A\, e^{-\Delta t/\tau}, & \Delta t > 0 \\ -A\, e^{\Delta t/\tau}, & \Delta t < 0 \end{cases}, \qquad \text{(Equation 4).}$$

Here, $A = 0.01$ and $\tau = 20\ ms$ exhibit the maximum magnitude and time constant of the STDP function for the synaptic potentiation or depression.

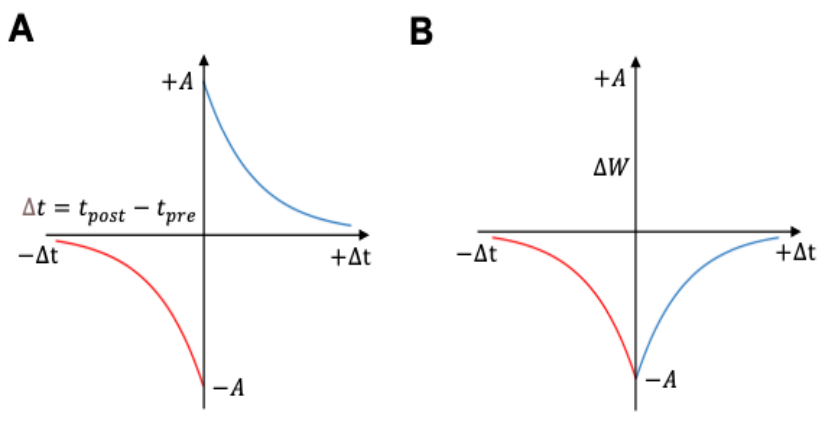


Figure 7. STPS curves. (A) Classical STDP curve showing relationship between synaptic weight change and the precise time difference between the Kenyon Cells and MBON spikes. The synaptic weight can be either depressed or potentiated. **(B)** STDP curve modulated by octopamine in the insect mushroom body. The Synaptic weights are depressed. The formula of these curves were described in Equations 3 and 4.

To train the model in different conditions of scanning, the flight-scan forms of the positive and negative patterns were presented to the model. Each set of flight-scan input contained a set of five patches with size 75x75 pixels were selected from the test patterns by shifting 15 pixels over each pattern from the left to right (Figure 2A). The numbers of shifted pixels control the speed of scanning. The activity of the MBON was used to assess the performance of the model. Following the training, the performance of the model was calculated from a decrease in firing rate of the MBON to a pattern that had been rewarding and/or an increase in firing rate of MBON to a pattern that had been punishing in training. The bee's final behavioural decision is proposed to come from a simple integration of these different valence-encoding neurons.

Acknowledgements

We thank Paul Graham and Andrew Barron for valuable comments on manuscript and Alice Bridges for drawing the front view of the bumblebee presented in Figure 1A. This research was financed by HFSP programme grant RGP0022/2014 and by EPSRC program grant Brains-on-Board EP/Poo6094/1.

Conflict of interest statement

All authors declare no conflict of interest.

Data and code accessibility

The code developed for this research project has been made openly accessible on GitHub: https://github.com/hadiimaboudi/neuromorphic_model_active-vision-

References

Abbott LF, Nelson SB. 2000. Synaptic plasticity: taming the beast. *Nat Neurosci* **3**:1178–1183.

Anderson AG, Ratnam K, Roorda A, Olshausen BA. 2020. High-acuity vision from retinal image motion. *J Vis* **20**:34–34.

Arenas A, Farina WM. 2008. Age and rearing environment interact in the retention of early olfactory memories in honeybees. *J Comp Physiol A* **194**:629–640. doi:10.1007/s00359-008-0337-z

Arenz A, Drews MS, Richter FG, Ammer G, Borst A. 2017. The temporal tuning of the Drosophila motion detectors is determined by the dynamics of their input elements. *Curr Biol* **27**:929–944.

Avarguès-Weber A, Dyer AG, Giurfa M. 2011. Conceptualization of above and below relationships by an insect. *Proc R Soc Lond B Biol Sci* **278**:898–905. doi:10.1098/rspb.2010.1891

Barlow HB. 1961. Possible principles underlying the transformation of sensory messages. *Sens Commun* **1**.

Benard J, Stach S, Giurfa M. 2006. Categorization of visual stimuli in the honeybee Apis mellifera. *Anim Cogn* **9**:257–270. doi:10.1007/s10071-006-0032-9

Boeddeker N, Mertes M, Dittmar L, Egelhaaf M. 2015. Bumblebee Homing: The Fine Structure of Head Turning Movements. *PLOS ONE* **10**:e0135020. doi:10.1371/journal.pone.0135020

Brinkworth RS, O'Carroll DC. 2009. Robust models for optic flow coding in natural scenes inspired by insect biology. *PLoS Comput Biol* **5**:e1000555.

Burke CJ, Huetteroth W, Owald D, Perisse E, Krashes MJ, Das G, Gohl D, Silies M, Certel S, Waddell S. 2012. Layered reward signalling through octopamine and dopamine in Drosophila. *Nature* **492**:433–437.

Caron SJC, Ruta V, Abbott LF, Axel R. 2013. Random convergence of olfactory inputs in the Drosophila mushroom body. *Nature* **497**:113–117. doi:10.1038/nature12063

Cassenaer S, Laurent G. 2012. Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature* **482**:47–52. doi:10.1038/nature10776

Cassenaer S, Laurent G. 2007. Hebbian STDP in mushroom bodies facilitates the synchronous flow of olfactory information in locusts. *Nature* **448**:709–713. doi:10.1038/nature05973

Chittka L, Niven J. 2009. Are bigger brains better? *Curr Biol CB* **19**:R995–R1008. doi:10.1016/j.cub.2009.08.023

Chittka L, Skorupski P. 2017. Active vision: a broader comparative perspective is needed.

Clark DA, Fitzgerald JE, Ales JM, Gohl DM, Silies MA, Norcia AM, Clandinin TR. 2014. Flies and humans share a motion estimation strategy that exploits natural scene statistics. *Nat Neurosci* **17**:296–303. doi:10.1038/nn.3600

Cognigni P, Felsenberg J, Waddell S. 2018. Do the right thing: neural network mechanisms of memory formation, expression and update in Drosophila. *Curr Opin Neurobiol* **49**:51–58. doi:10.1016/j.conb.2017.12.002

Collett TS, Fry SN, Wehner R. 1993. Sequence learning by honeybees. *J Comp Physiol A* **172**:693–706. doi:10.1007/BF00195395

Cope AJ, Vasilaki E, Minors D, Sabo C, Marshall JAR, Barron AB. 2018. Abstract concept learning in a simple neural network inspired by the insect brain. *PLOS Comput Biol* **14**:e1006435. doi:10.1371/journal.pcbi.1006435

Dawkins MS, Woodington A. 2000. Pattern recognition and active vision in chickens. *Nature* **403**:652–655. doi:10.1038/35001064

de Croon GCHE, Dupeyroux JJG, Fuller SB, Marshall JAR. 2022. Insect-inspired AI for autonomous robots. *Sci Robot* **7**:eabl6334. doi:10.1126/scirobotics.abl6334

Doussot C, Bertrand OJN, Egelhaaf M. 2021. The Critical Role of Head Movements for Spatial Representation During Bumblebees Learning Flight. *Front Behav Neurosci* **14**.

Dyakova O, Lee Y-J, Longden KD, Kiselev VG, Nordström K. 2015. A higher order visual neuron tuned to the spatial amplitude spectra of natural scenes. *Nat Commun* **6**:8522. doi:10.1038/ncomms9522

Dyakova O, Müller MM, Egelhaaf M, Nordström K. 2019. Image statistics of the environment surrounding freely behaving hoverflies. *J Comp Physiol A* **205**:373–385.

Dyakova O, Nordström K. 2017. Image statistics and their processing in insect vision. *Curr Opin Insect Sci*, Neuroscience * Pheromones **24**:7–14. doi:10.1016/j.cois.2017.08.002

Dyer AG, Neumeyer C, Chittka L. 2005. Honeybee (Apis mellifera) vision can discriminate between and recognise images of human faces. *J Exp Biol* **208**:4709–4714.

Dyer AG, Paulk AC, Reser DH. 2011. Colour processing in complex environments: insights from the visual system of bees. *Proc R Soc Lond B Biol Sci* **278**:952–959. doi:10.1098/rspb.2010.2412

Fenk LM, Avritzer SC, Weisman JL, Nair A, Randt LD, Mohren TL, Siwanowicz I, Maimon G. 2022. Muscles that move the retina augment compound eye vision in Drosophila. *Nature* **612**:116–122. doi:10.1038/s41586-022-05317-5

Fisher YE, Silies M, Clandinin TR. 2015. Orientation Selectivity Sharpens Motion Detection in Drosophila. *Neuron* **88**:390–402. doi:10.1016/j.neuron.2015.09.033

Geisler WS. 2008. Visual Perception and the Statistical Properties of Natural Scenes. *Annu Rev Psychol* **59**:167–192. doi:10.1146/annurev.psych.58.110405.085632

Gibson JJ. 1979. The ecological approach to visual perception.

Giger AD, Srinivasan MV. 1996. Pattern recognition in honeybees: chromatic properties of orientation analysis. *J Comp Physiol A* **178**:763–769. doi:10.1007/BF00225824

Giurfa M. 2013. Cognition with few neurons: higher-order learning in insects. *Trends Neurosci* **36**:285–294. doi:10.1016/j.tins.2012.12.011

Giurfa M, Hammer M, Stach S, Stollhoff N, MüLler-Deisig N, Mizyrycki C. 1999. Pattern learning by honeybees: conditioning procedure and recognition strategy. *Anim Behav* **57**:315–324. doi:10.1006/anbe.1998.0957

Giurfa M, Zhang S, Jenett A, Menzel R, Srinivasan MV. 2001. The concepts of 'sameness' and 'difference' in an insect. *Nature* **410**:930–933. doi:10.1038/35073582

Gribakin FG. 1975. Functional morphology of the compound eye of the bee. *Compd Eye Vis Insects* **154**:176.

Guiraud M, Roper M, Chittka L. 2018. High-Speed Videography Reveals How Honeybees Can Turn a Spatial Concept Learning Task Into a Simple Discrimination Task by Stereotyped Flight Movements and Sequential Inspection of Pattern Elements. *Front Psychol* **9**. doi:10.3389/fpsyg.2018.01347

Hammer M. 1993. An identified neuron mediates the unconditioned stimulus in associative olfactory learning in honeybees. *Nature* **366**:59–63. doi:10.1038/366059a0

Hammer M, Menzel R. 1995. Learning and memory in the honeybee. *J Neurosci* **15**:1617–1630.

Hateren J van, Srinivasan MV, Wait PB. 1990. Pattern recognition in bees: orientation discrimination. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* **167**:649–654.

Heisenberg M. 2003. Mushroom body memoir: from maps to models. *Nat Rev Neurosci* **4**:266–275. doi:10.1038/nrn1074

Hertel H. 1983. Change of synapse frequency in certain photoreceptors of the honeybee after chromatic deprivation. *J Comp Physiol* **151**:477–482.

Hertel H. 1982. The effect of spectral light deprivation on the spectral sensitivity of the honey bee. *J Comp Physiol* **147**:365–369.

Honegger KS, Campbell RA, Turner GC. 2011. Cellular-resolution population imaging reveals robust sparse coding in the Drosophila mushroom body. *J Neurosci* **31**:11772–11785.

Hyvärinen A, Hurri J, Hoyer PO. 2009. Natural image statistics: A probabilistic approach to early computational vision. Springer Science & Business Media.

James AC, Osorio D. 1996. Characterisation of columnar neurons and visual signal processing in the medulla of the locust optic lobe by system identification techniques. *J Comp Physiol A* **178**:183–199. doi:10.1007/BF00188161

Juusola M, Dau A, Song Z, Solanki N, Rien D, Jaciuch D, Dongre SA, Blanchard F, Polavieja GG de, Hardie RC, Takalo J. 2017. Microsaccadic sampling of moving image information provides Drosophila hyperacute vision. *eLife* **6**:e26117. doi:10.7554/eLife.26117

Juusola M, Song Z. 2017. How a fly photoreceptor samples light information in time. *J Physiol* **595**:5427–5437.

Kuang X, Poletti M, Victor JD, Rucci M. 2012. Temporal encoding of spatial information during active visual fixation. *Curr Biol CB* **22**:510–514. doi:10.1016/j.cub.2012.01.050

Land MF. 1999. Motion and vision: why animals move their eyes. *J Comp Physiol [A]* **185**:341–352.

Land MF. 1973. Head Movement of Flies during Visually Guided Flight. *Nature* **243**:299–300. doi:10.1038/243299a0

Land MF, Nilsson D-E. 2012. Animal eyes. Oxford University Press.

Langridge KV, Wilke C, Riabinina O, Vorobyev M, de Ibarra NH. 2018. Approach direction prior to landing explains patterns of colour learning. *bioRxiv* 381210.

Langridge KV, Wilke C, Riabinina O, Vorobyev M, Hempel de Ibarra N. 2021. Approach Direction Prior to Landing Explains Patterns of Colour Learning in Bees. *Front Physiol* **12**.

Lehrer M, Collett TS. 1994. Approaching and departing bees learn different cues to the distance of a landmark. *J Comp Physiol A* **175**:171–177. doi:10.1007/BF00215113

Locatelli FF, Fernandez PC, Villareal F, Muezzinoglu K, Huerta R, Galizia CG, Smith BH. 2013. Nonassociative plasticity alters competitive interactions among mixture components in early olfactory processing. *Eur J Neurosci* **37**:63–79. doi:10.1111/ejn.12021

MaBouDi H, Barron AB, Li S, Honkanen M, Loukola OJ, Peng F, Li W, Marshall JAR, Cope A, Vasilaki E, Solvi C. 2021a. Non-numerical strategies used by bees to solve numerical cognition tasks. *Proc R Soc B Biol Sci* **288**:20202711. doi:10.1098/rspb.2020.2711

MaBouDi H, Dona H, Gatto E, Loukola OJ, Buckley E, Onoufriou PD, Skorupski P, Chittka L. 2020a. Bumblebees use sequential scanning of countable items in visual patterns to solve numerosity tasks. *Integr Comp Biol*.

MaBouDi H, Marshall JA, Dearden N, Barron AB. 2023. How honey bees make fast and accurate decisions. *bioRxiv* 2023–01.

MaBouDi H, Marshall JAR, Barron AB. 2020b. Honeybees solve a multi-comparison ranking task by probability matching. *Proc R Soc B Biol Sci* **287**:20201525. doi:10.1098/rspb.2020.1525

MaBouDi H, Roper M, Guiraud M, Marshall JAR, Chittka L. 2021b. Automated video tracking and flight analysis show how bumblebees solve a pattern discrimination task using active vision. *bioRxiv* 2021.03.09.434580. doi:10.1101/2021.03.09.434580

MaBouDi H, Shimazaki H, Giurfa M, Chittka L. 2017. Olfactory learning without the mushroom bodies: Spiking neural network models of the honeybee lateral antennal lobe tract reveal its capacities in odour memory tasks of varied complexities. *PLoS Comput Biol* **13**:e1005551.

MaBouDi H, Solvi C, Chittka L. 2020c. Bumblebees learn a relational rule but switch to a win-stay/lose-switch heuristic after extensive training. *bioRxiv* 2020.05.08.085142. doi:10.1101/2020.05.08.085142

Markram H, Lübke J, Frotscher M, Sakmann B. 1997. Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science* **275**:213–215.

Matthews T, Osorio D, Cavallaro A, Chittka L. 2018. The importance of spatial visual scene parameters in predicting optimal cone sensitivities in routinely trichromatic frugivorous Old-World primates. *Front Comput Neurosci* **12**:15.

Meeks JP, Holy TE. 2008. Pavlov's moth: olfactory learning and spike timing–dependent plasticity. *Nat Neurosci* **11**:1126–1127.

Menzel R. 2012. The honeybee as a model for understanding the basis of cognition. *Nat Rev Neurosci* **13**:758–768. doi:10.1038/nrn3357

Menzel R, Blakers M. 1976. Colour receptors in the bee eye — Morphology and spectral sensitivity. *J Comp Physiol* **108**:11–13. doi:10.1007/BF00625437

Menzel R, Giurfa M. 2006. Dimensions of cognition in an insect, the honeybee. *Behav Cogn Neurosci Rev* **5**:24–40. doi:10.1177/1534582306289522

Nityananda V, Skorupski P, Chittka L. 2014. Can bees see at a glance? *J Exp Biol* **217**:1933–1939. doi:10.1242/jeb.101394

Odenthal L, Doussot C, Meyer S, Bertrand OJN. 2021. Analysing Head-Thorax Choreography During Free-Flights in Bumblebees. *Front Behav Neurosci* **14**. doi:10.3389/fnbeh.2020.610029

Oja E. 1982. Simplified neuron model as a principal component analyzer. *J Math Biol* **15**:267–273.

Okada R, Rybak J, Manz G, Menzel R. 2007. Learning-Related Plasticity in PE1 and Other Mushroom Body-Extrinsic Neurons in the Honeybee Brain. *J Neurosci* **27**:11736–11747. doi:10.1523/JNEUROSCI.2216-07.2007

Olshausen BA. 2003. Principles of image representation in visual cortex. *Vis Neurosci* **2**:1603–1615.

Paulk AC, Dacks AM, Phillips-Portillo J, Fellous J-M, Gronenberg W. 2009. Visual Processing in the Central Bee Brain. *J Neurosci* **29**:9987–9999. doi:10.1523/JNEUROSCI.1325-09.2009

Paulk AC, Gronenberg W. 2008. Higher order visual input to the mushroom bodies in the bee, Bombus impatiens. *Arthropod Struct Dev* **37**:443–458. doi:10.1016/j.asd.2008.03.002

Paulk AC, Phillips-Portillo J, Dacks AM, Fellous J-M, Gronenberg W. 2008. The Processing of Color, Motion, and Stimulus Timing Are Anatomically Segregated in the Bumblebee Brain. *J Neurosci* **28**:6319–6332. doi:10.1523/JNEUROSCI.1196-08.2008

Perry CJ, Barron AB. 2013. Neural mechanisms of reward in insects. *Annu Rev Entomol* **58**:543–562.

Roper M, Fernando C, Chittka L. 2017. Insect Bio-inspired Neural Network Provides New Evidence on How Simple Feature Detectors Can Enable Complex Visual Generalization and Stimulus Location Invariance in the Miniature Brain of Honeybees. *PLOS Comput Biol* **13**:e1005333. doi:10.1371/journal.pcbi.1005333

Rucci M, Iovin R, Poletti M, Santini F. 2007. Miniature eye movements enhance fine spatial detail. *Nature* **447**:852–855.

Rucci M, Victor JD. 2015. The unsteady eye: an information-processing stage, not a bug. *Trends Neurosci* **38**:195–206. doi:10.1016/j.tins.2015.01.005

Ruderman DL. 1994. The statistics of natural images. *Netw Comput Neural Syst* **5**:517.

Schwaerzel M, Monastirioti M, Scholz H, Friggi-Grelin F, Birman S, Heisenberg M. 2003. Dopamine and octopamine differentiate between aversive and appetitive olfactory memories in Drosophila. *J Neurosci Off J Soc Neurosci* **23**:10495–10502.

Schwegmann A, Lindemann JP, Egelhaaf M. 2014. Temporal statistics of natural image sequences generated by movements with insect flight characteristics. *PloS One* **9**:e110386. doi:10.1371/journal.pone.0110386

Seelig JD, Jayaraman V. 2013. Feature detection and orientation tuning in the Drosophila central complex. *Nature* **503**:262–266. doi:10.1038/nature12601

Serbe E, Meier M, Leonhardt A, Borst A. 2016. Comprehensive characterization of the major presynaptic elements to the Drosophila OFF motion detector. *Neuron* **89**:829–841.

Simoncelli EP, Olshausen BA. 2001. Natural image statistics and neural representation. *Annu Rev Neurosci* **24**:1193–1216.

Skorupski P, Döring TF, Chittka L. 2007. Photoreceptor spectral sensitivity in island and mainland populations of the bumblebee, Bombus terrestris. *J Comp Physiol A* **193**:485–494.

Song S, Miller KD, Abbott LF. 2000. Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nat Neurosci* **3**:919–926.

Song Z, Juusola M. 2014. Refractory Sampling Links Efficiency and Costs of Sensory Encoding to Stimulus Statistics. *J Neurosci* **34**:7216–7237. doi:10.1523/JNEUROSCI.4463-13.2014

Spaethe J, Chittka L. 2003. Interindividual variation of eye optics and single object resolution in bumblebees. *J Exp Biol* **206**:3447–3453.

Spaethe J, Tautz J, Chittka L. 2001. Visual constraints in foraging bumblebees: flower size and color affect search time and flight behavior. *Proc Natl Acad Sci* **98**:3898–3903.

Srinivasan MV. 2010. Honey bees as a model for vision, perception, and cognition. *Annu Rev Entomol* **55**:267–284.

Srinivasan MV. 1994. Pattern recognition in the honeybee: recent progress. *J Insect Physiol* **40**:183–194.

Srinivasan MV, Lehrer M. 1988. Spatial acuity of honeybee vision and its spectral properties. *J Comp Physiol A* **162**:159–172. doi:10.1007/BF00606081

Stach S, Benard J, Giurfa M. 2004. Local-feature assembling in visual pattern recognition and generalization in honeybees. *Nature* **429**:758–761. doi:10.1038/nature02594

Stach S, Giurfa M. 2005. The influence of training length on generalization of visual feature assemblies in honeybees. *Behav Brain Res* **161**:8–17. doi:10.1016/j.bbr.2005.02.008

Streinzer M, Brockmann A, Nagaraja N, Spaethe J. 2013. Sex and Caste-Specific Variation in Compound Eye Morphology of Five Honeybee Species. *PLOS ONE* **8**:e57702. doi:10.1371/journal.pone.0057702

Strube-Bloss MF, Nawrot MP, Menzel R. 2011. Mushroom Body Output Neurons Encode Odor–Reward Associations. *J Neurosci* **31**:3129–3140. doi:10.1523/JNEUROSCI.2583-10.2011

Szyszka P, Ditzen M, Galkin A, Galizia CG, Menzel R. 2005. Sparsening and temporal sharpening of olfactory representations in the honeybee mushroom bodies. *J Neurophysiol* **94**:3303–3313. doi:10.1152/jn.00397.2005

Taylor GJ, Tichit P, Schmidt MD, Bodey AJ, Rau C, Baird E. 2019. Bumblebee visual allometry results in locally improved resolution and globally improved sensitivity. *Elife* **8**:e40613.

Turner CH. 1911. Experiments on Pattern-Vision of the Honey Bee. *Biol Bull* **21**:249–264. doi:10.2307/1536017

Van Hateren JH. 1997. Processing of natural time series of intensities by the visual system of the blowfly. *Vision Res* **37**:3407–3416.

van Hateren JH. 1992. Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation. *J Comp Physiol A* **171**:157–170.

Vasas V, Peng F, MaBouDi H, Chittka L. 2019. Randomly weighted receptor inputs can explain the large diversity of colour-coding neurons in the bee visual system. *Sci Rep* **9**:1–13.

Vetter RS, Visscher PK. n.d. Influence of Age on Antennal Response of Male Honey Bees, Apis mellifera, to Queen Mandibular Pheromone and Alarm Pheromone Component. *J Chem Ecol* **23**:1867–1880. doi:10.1023/B:JOEC.0000006456.90528.94

Vogels TP, Sprekeler H, Zenke F, Clopath C, Gerstner W. 2011. Inhibitory Plasticity Balances Excitation and Inhibition in Sensory Pathways and Memory Networks. *Science* **334**:1569–1573. doi:10.1126/science.1211095

Von Frisch K. 1914. Der farbensinn und formensinn der biene. Рипол Классик.

Webb B. 2020. Robots with insect brains. *Science* **368**:244–245.

Wehner R. 1967. Pattern recognition in bees. *Nature* **215**:1244–1248.

Werner A, Stürzl W, Zanker J. 2016. Object Recognition in Flight: How Do Bees Distinguish between 3D Shapes? *PLoS ONE* **11**:e0147106. doi:10.1371/journal.pone.0147106

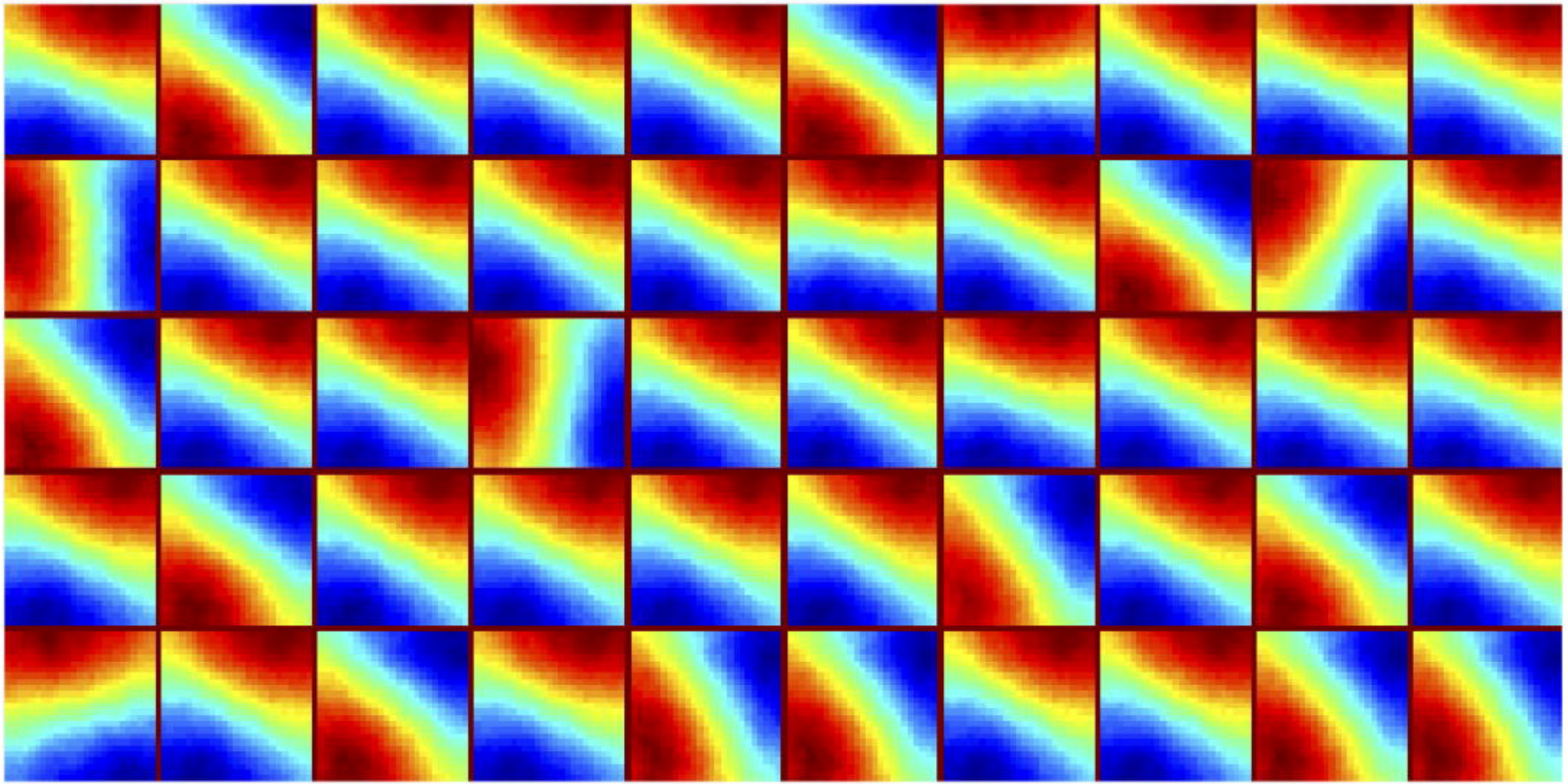
Yang E-C, Maddess T. 1997. Orientation-sensitive Neurons in the Brain of the Honey Bee (Apis mellifera). *J Insect Physiol* **43**:329–336. doi:10.1016/S0022-1910(96)00111-4

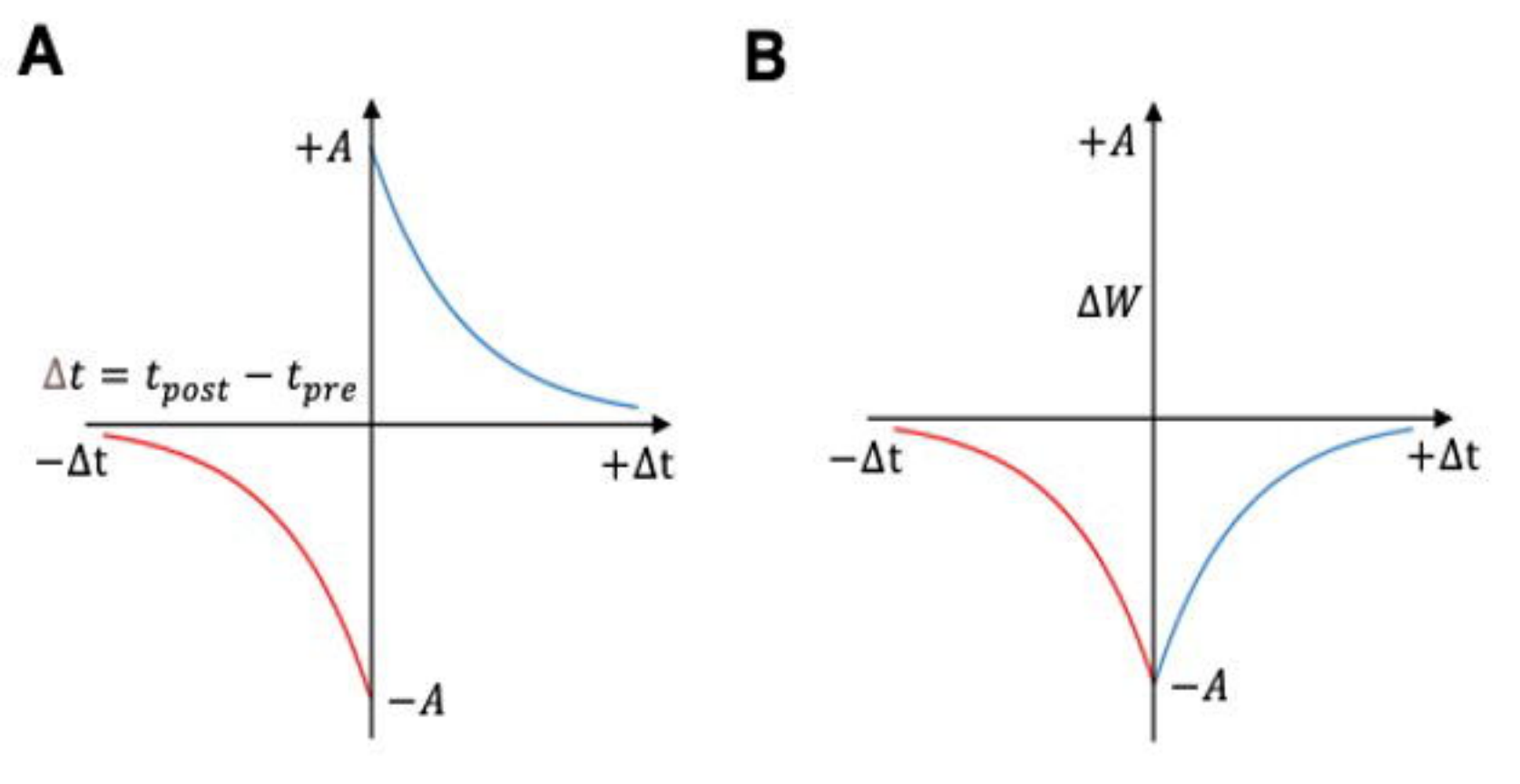
Yarbus AL. 2013. Eye movements and vision. Springer.

Zhang LI, Tao HW, Holt CE, Harris WA, Poo M. 1998. A critical window for cooperation and competition among developing retinotectal synapses. *Nature* **395**:37–44.

Zhang SW, Horridge GA. 1992. Pattern recognition in bees: size of regions in spatial layout. *Philos Trans R Soc Lond B Biol Sci* **337**:65–71.

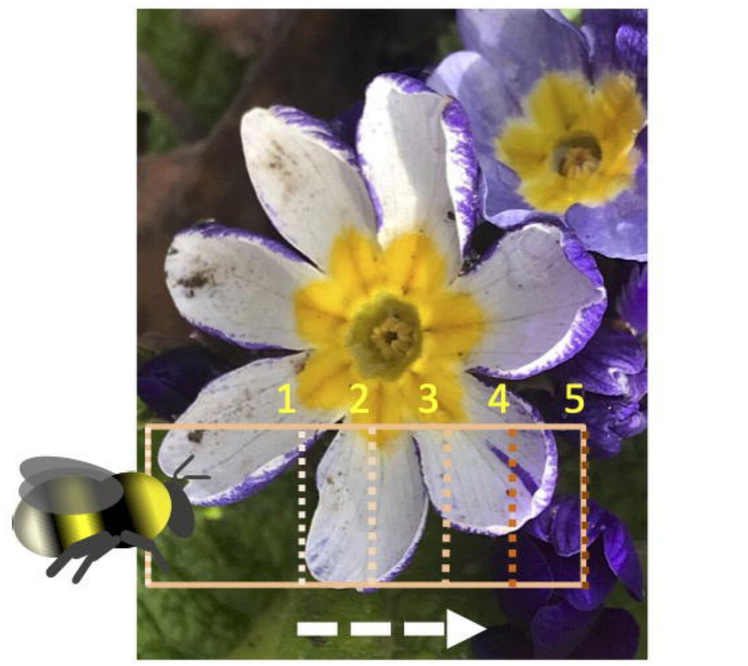
Zimmermann MJY, Nevala NE, Yoshimatsu T, Osorio D, Nilsson D-E, Berens P, Baden T. 2018. Zebrafish Differentially Process Color across Visual Space to Match Natural Scenes. *Curr Biol* **0**. doi:10.1016/j.cub.2018.04.075


-1
0
1
synaptic weights


A
+A
Δt = t_post − t_pre
−Δt
+Δt
−A
B
+A
ΔW
−Δt
+Δt
−A

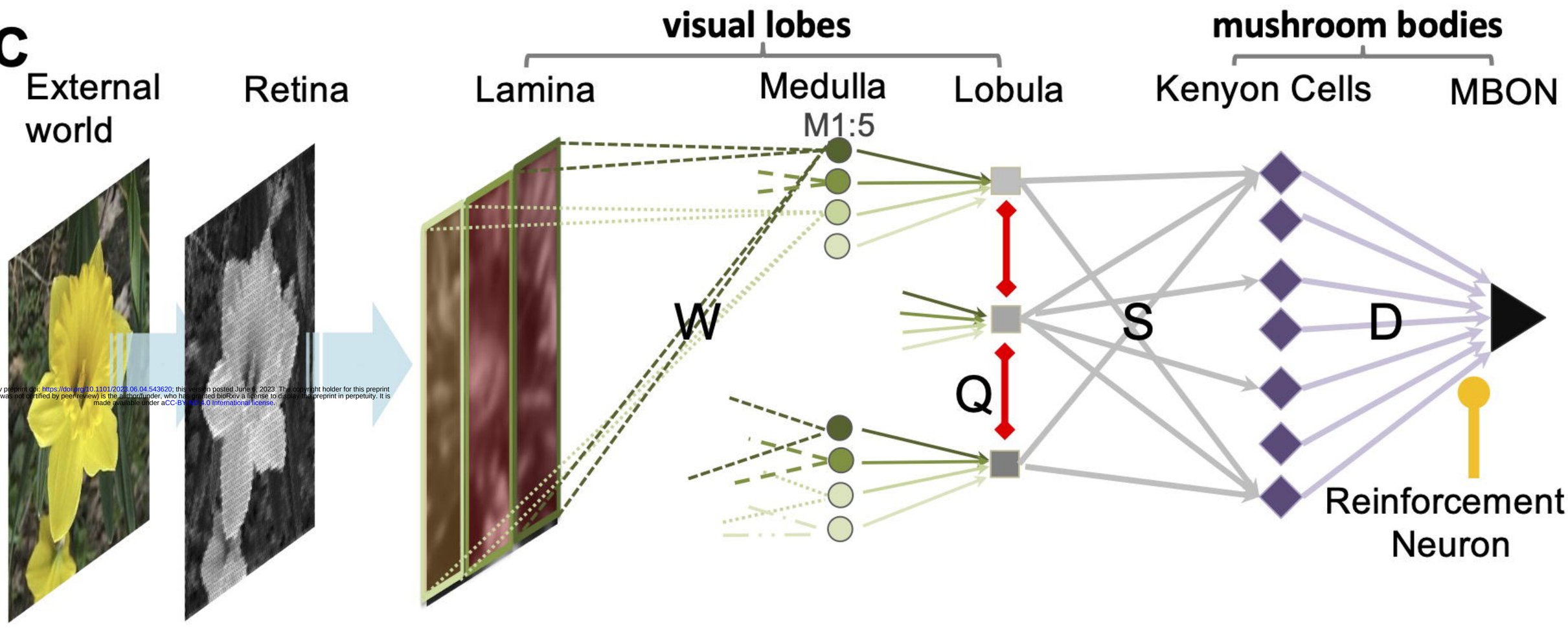
A

B

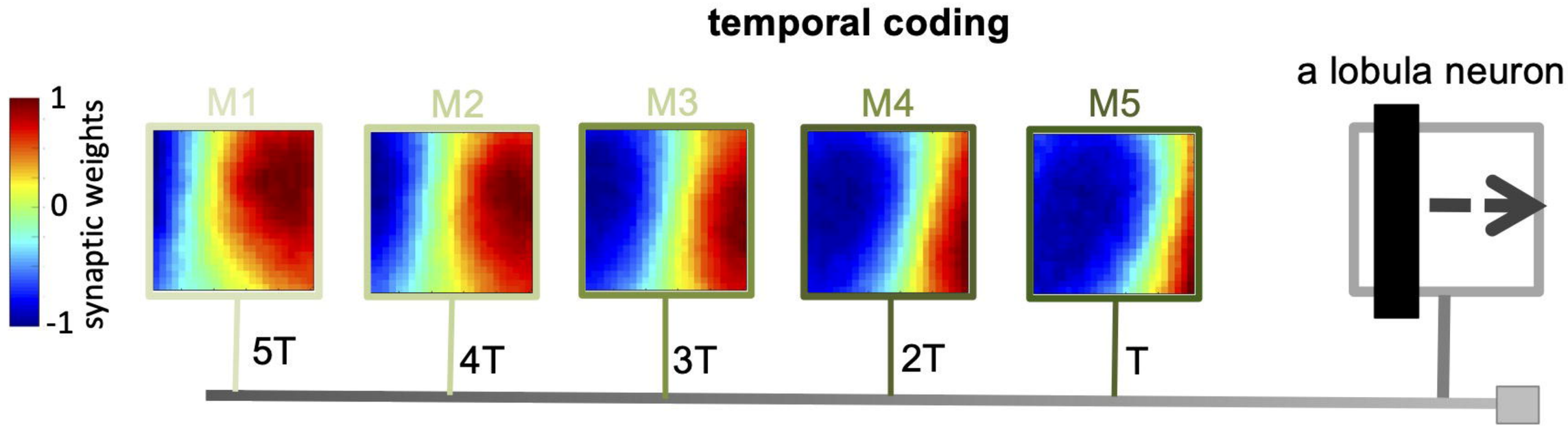


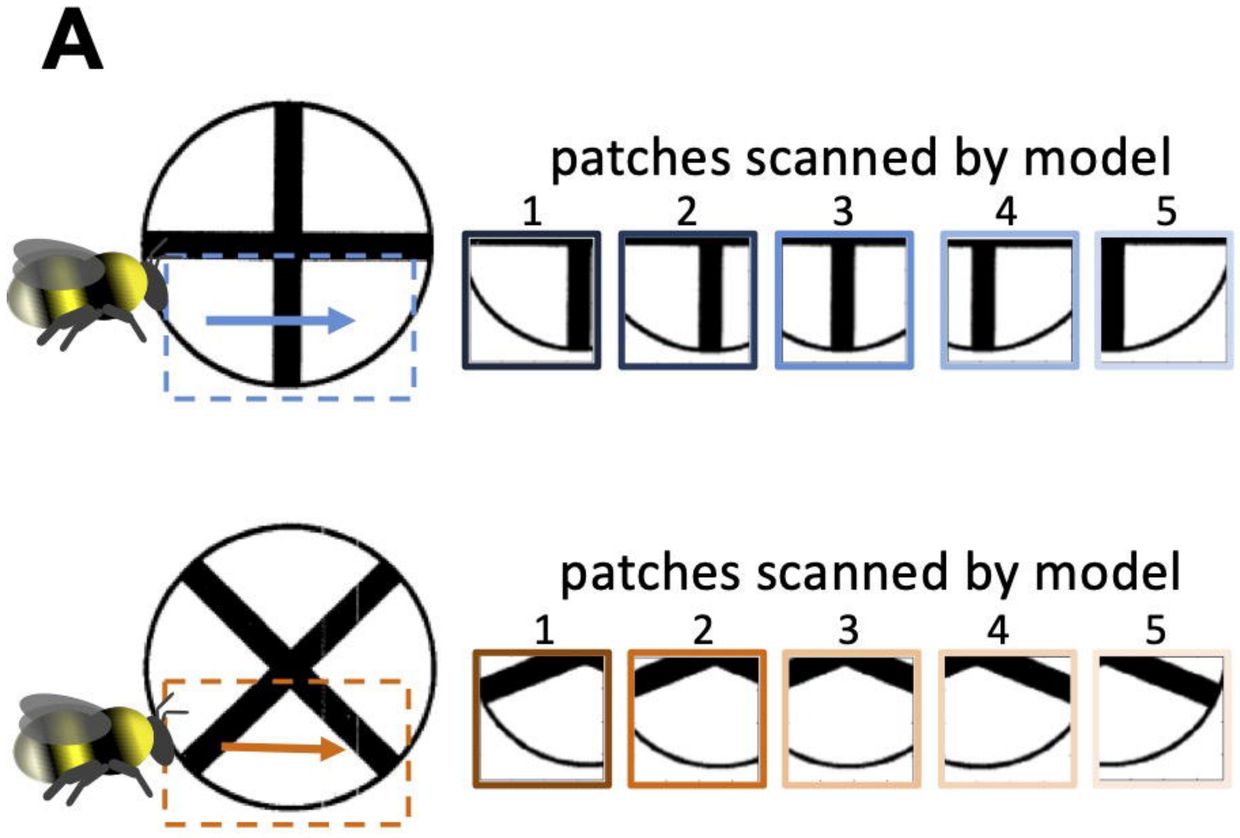
patches scanned by model

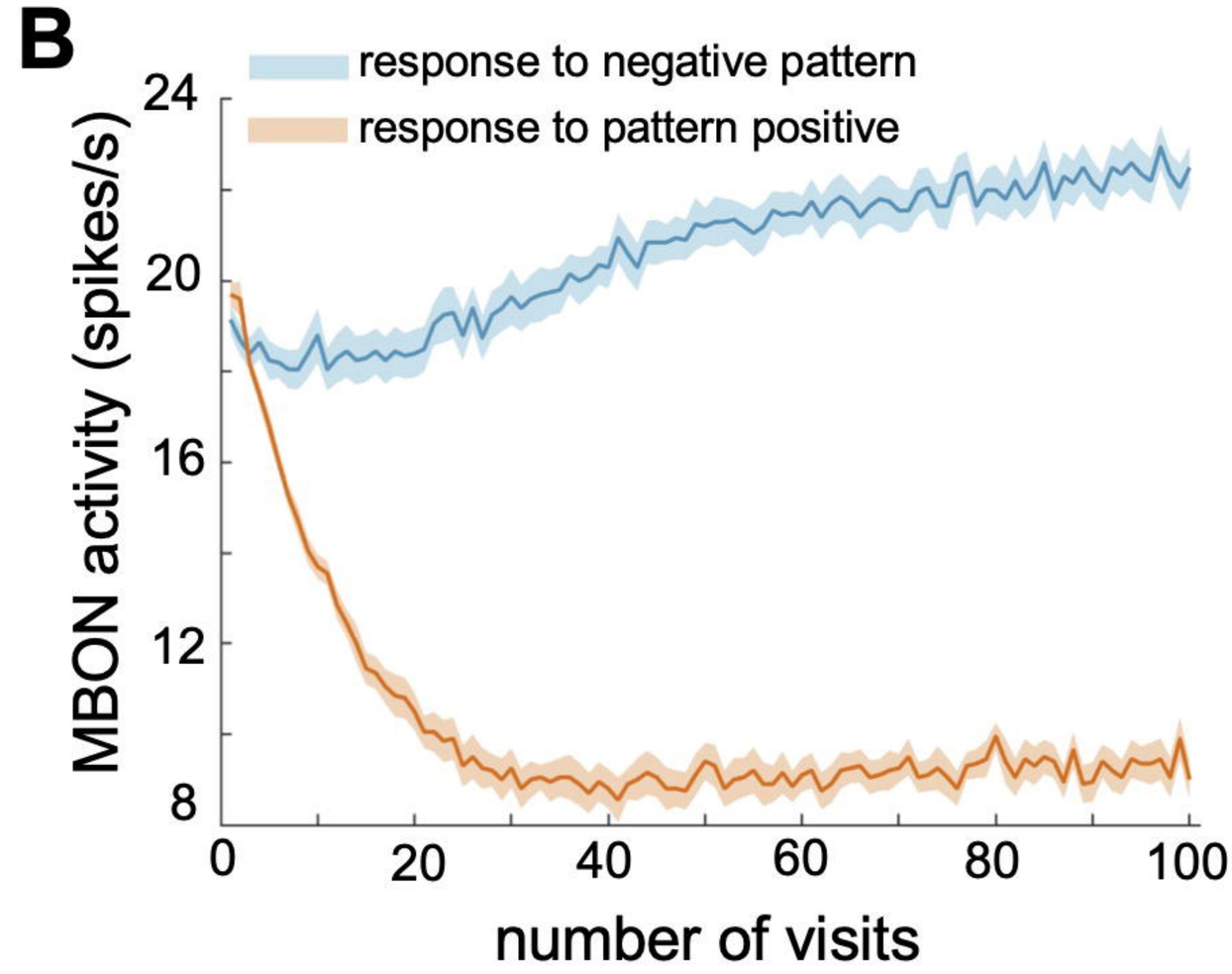
C

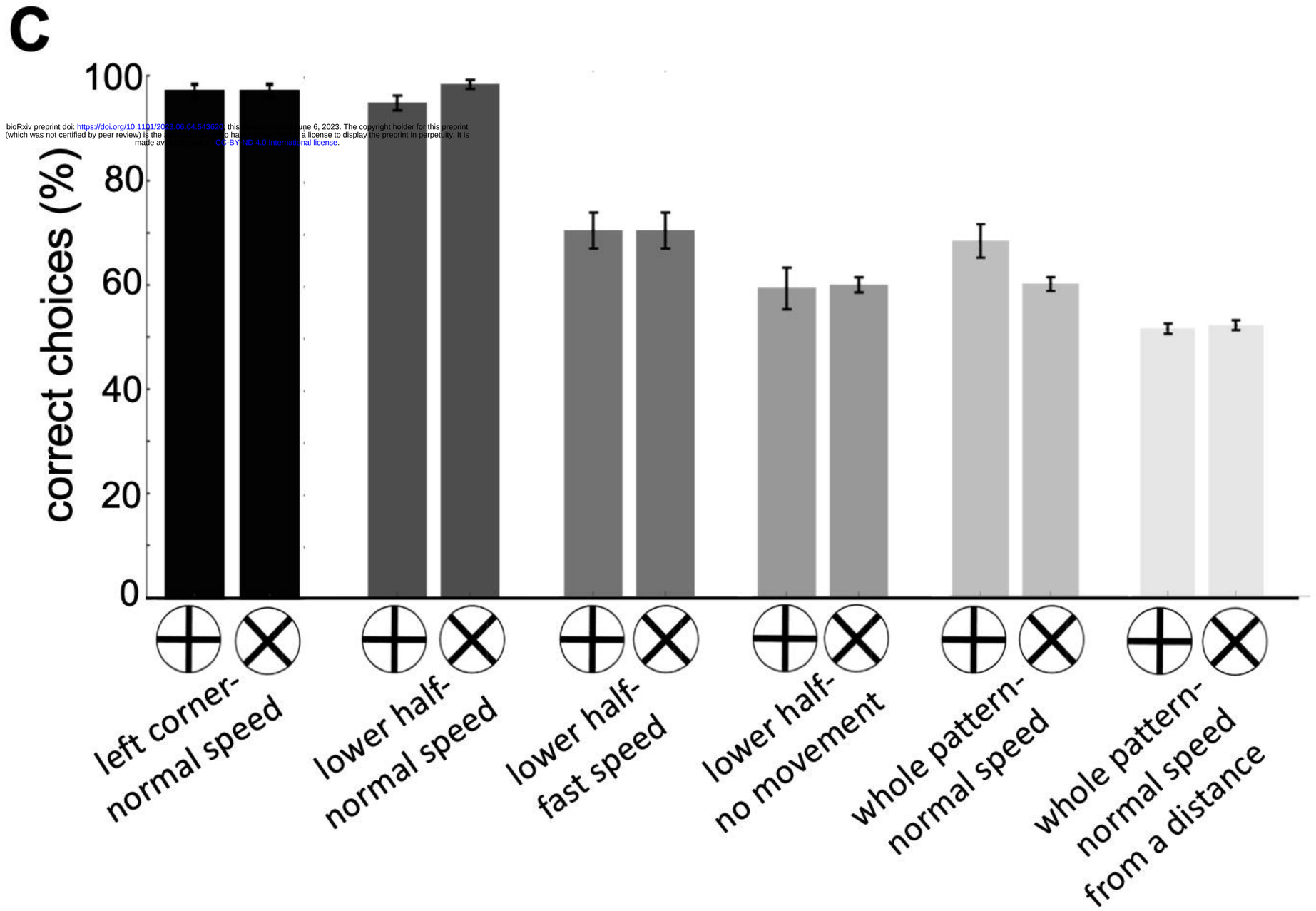


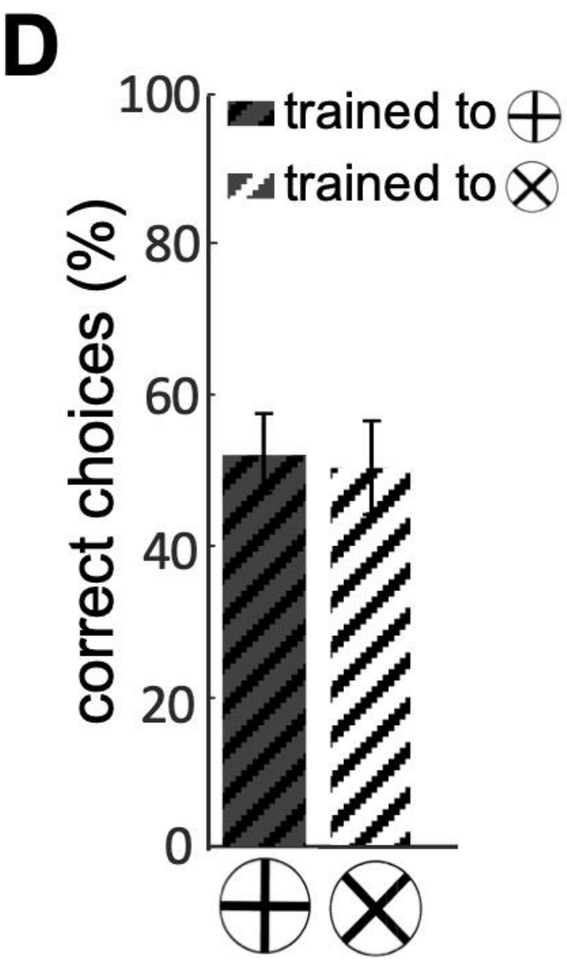
D

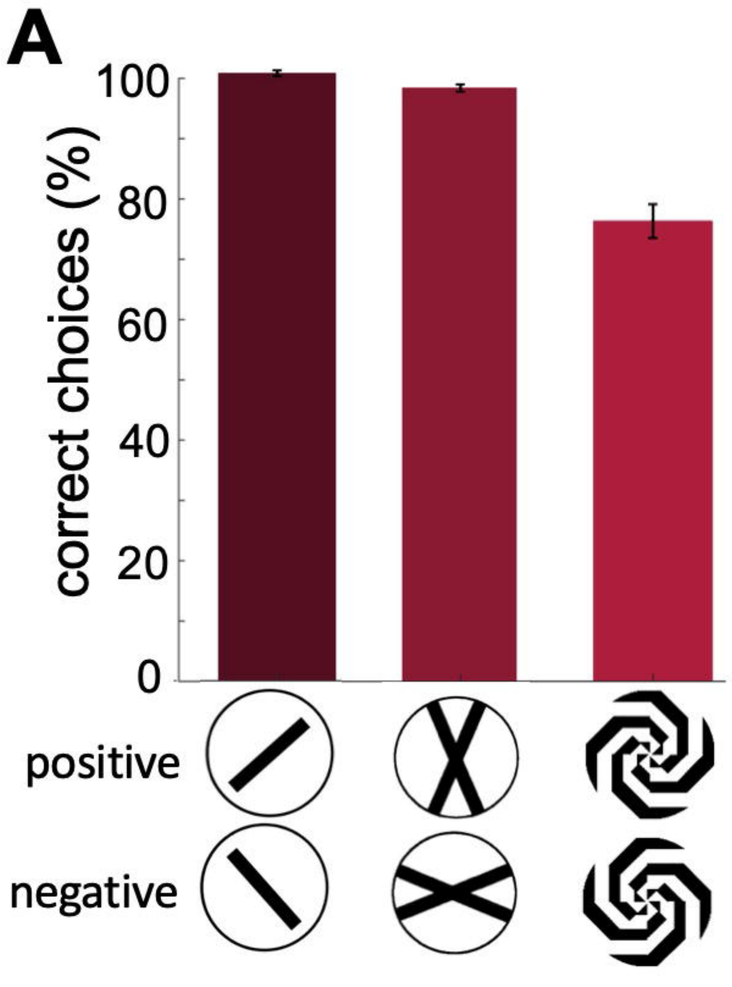


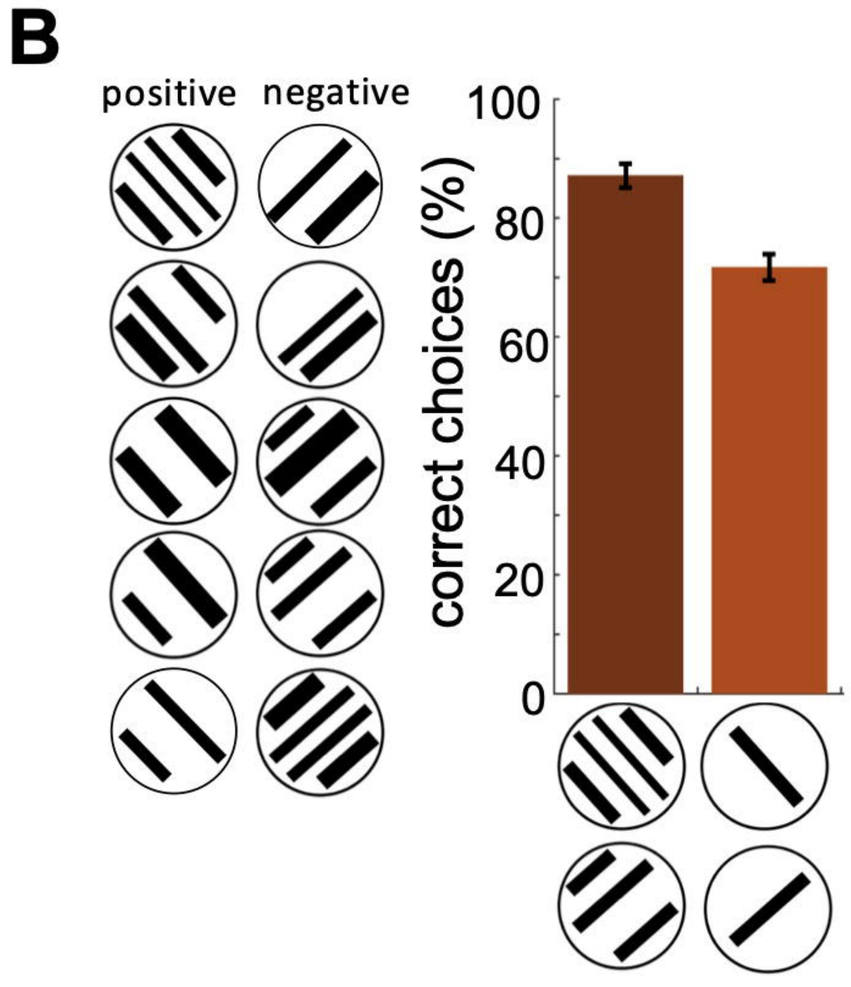


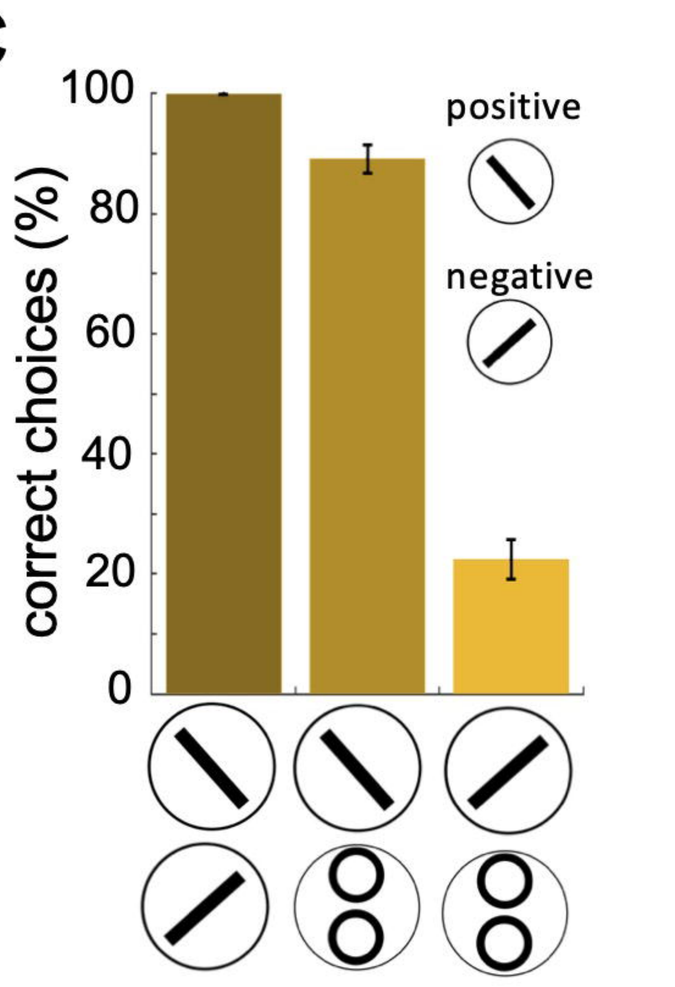


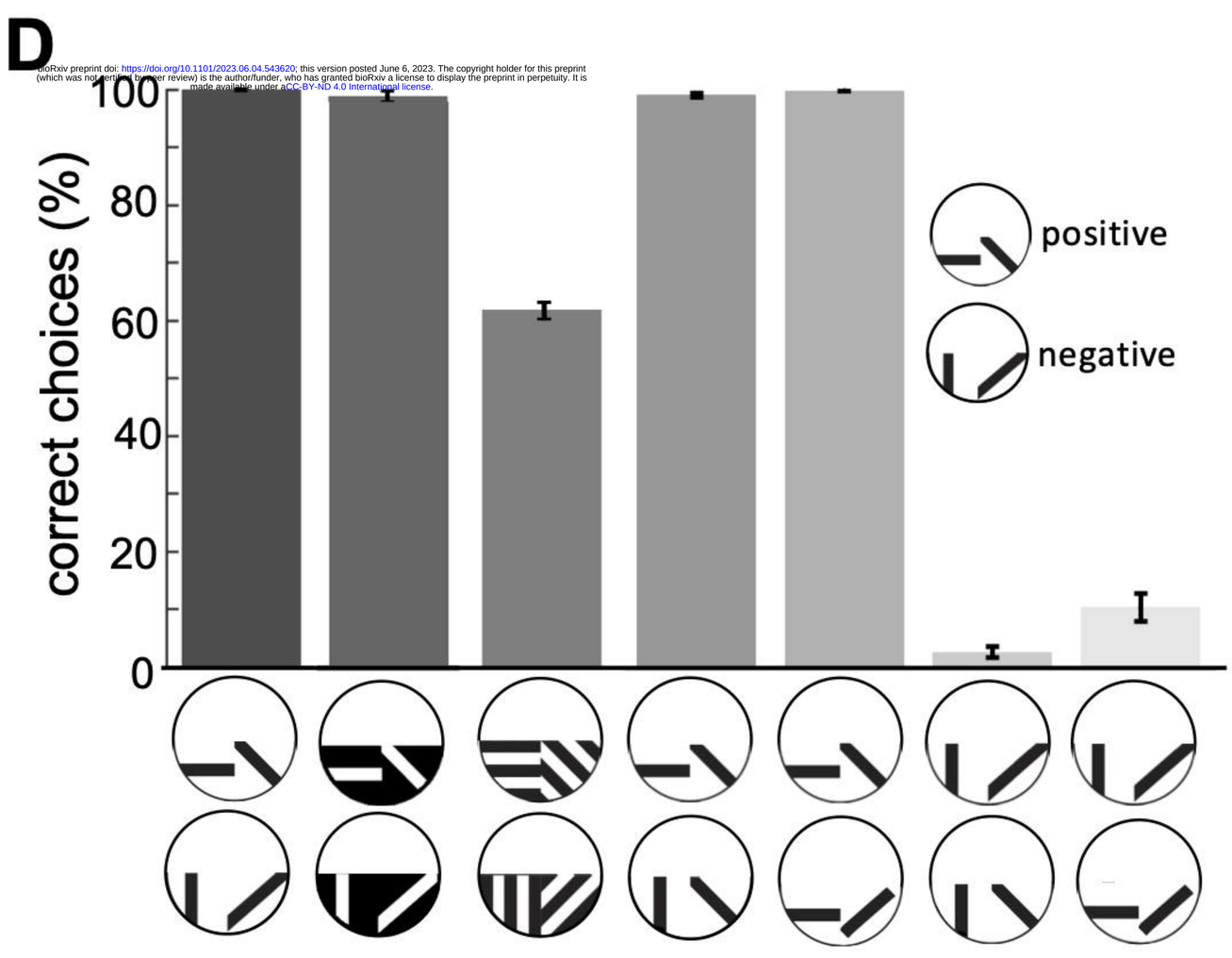


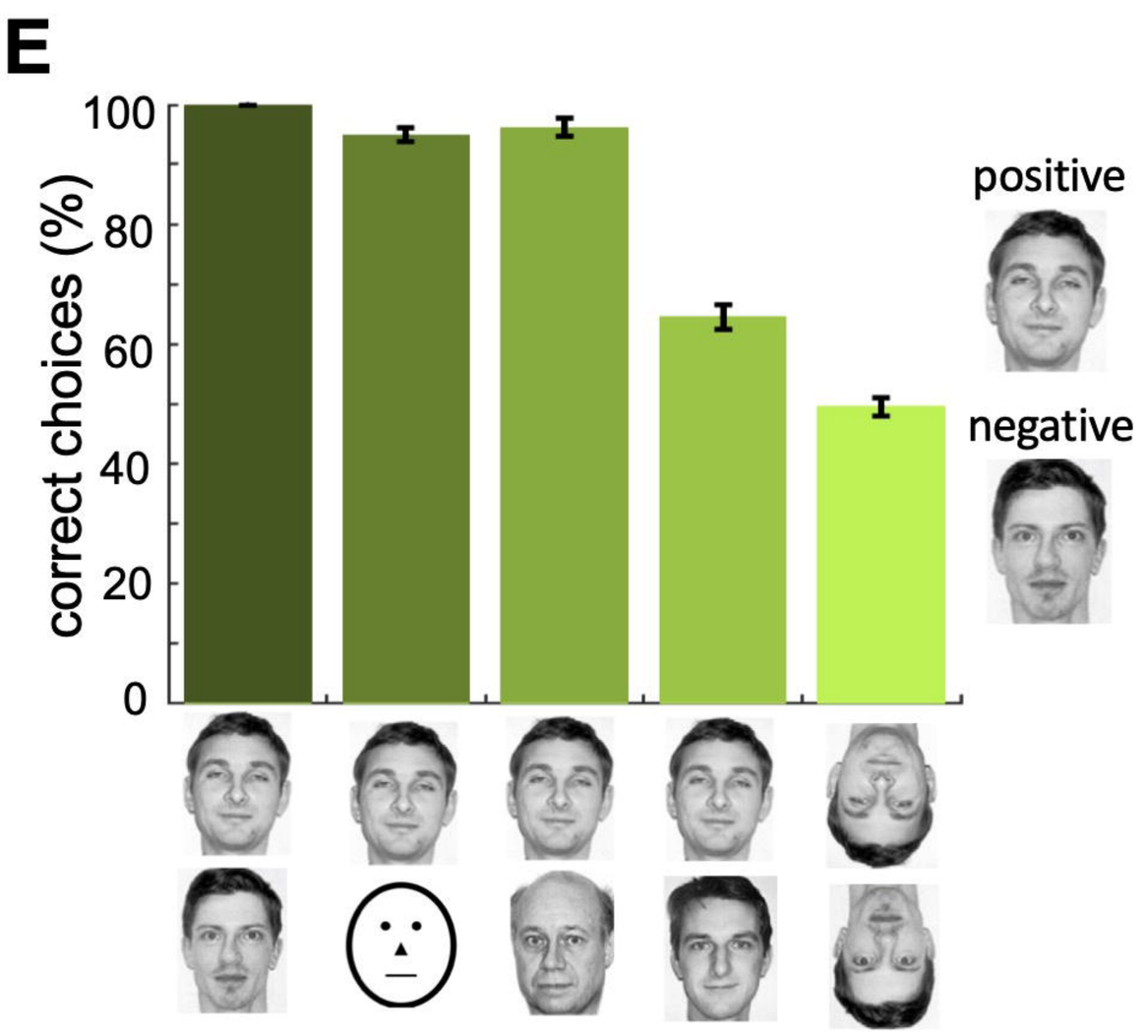


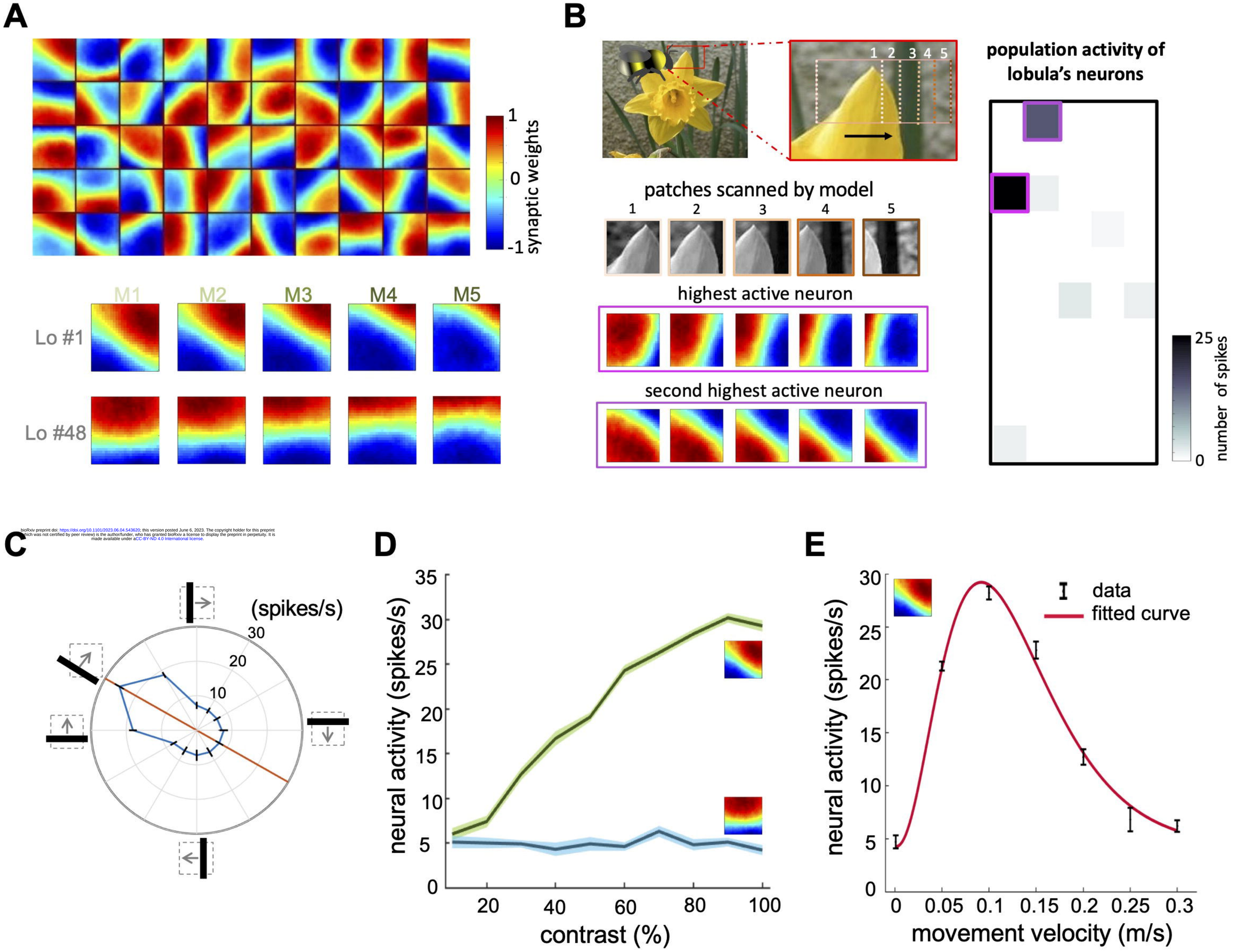






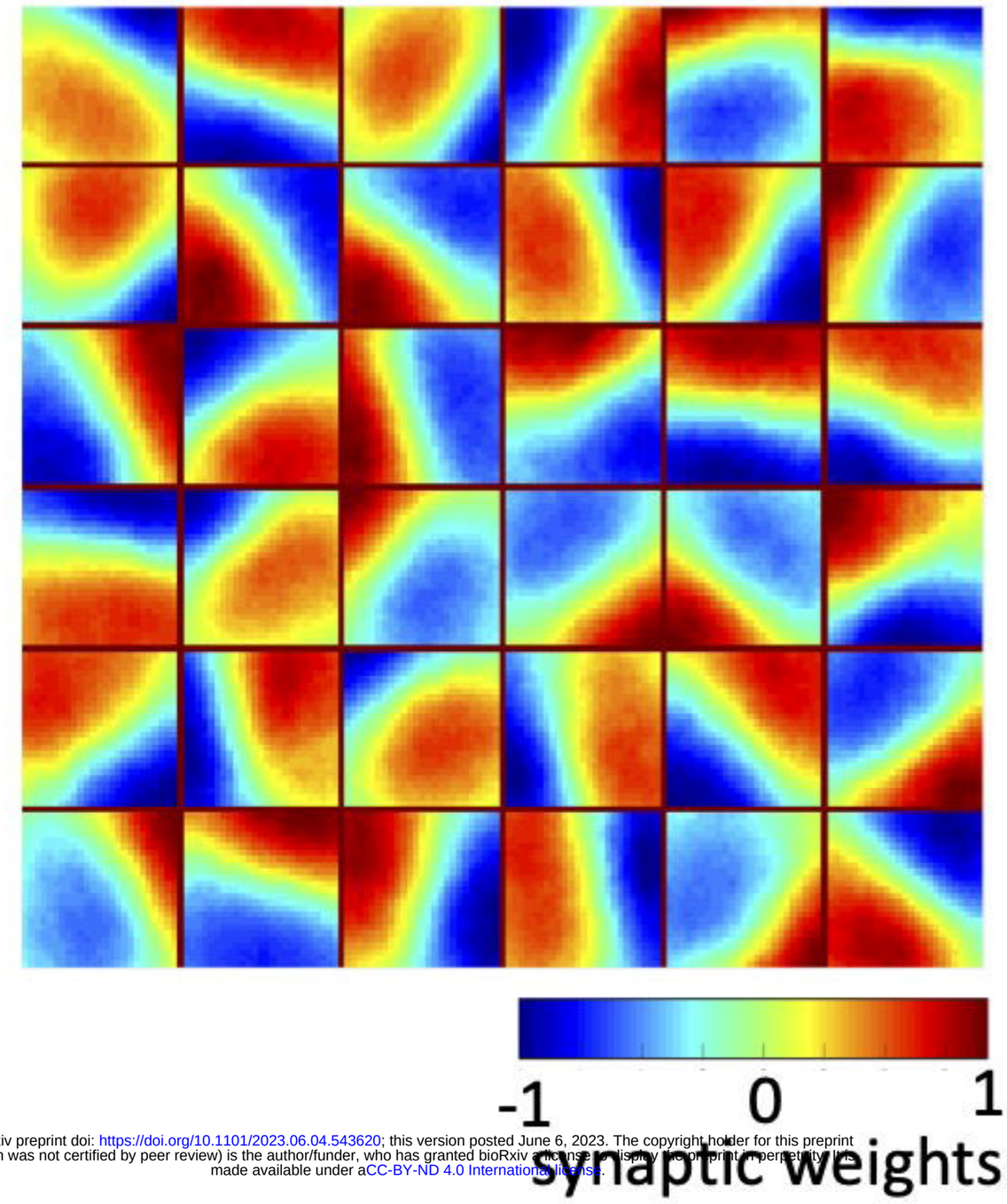




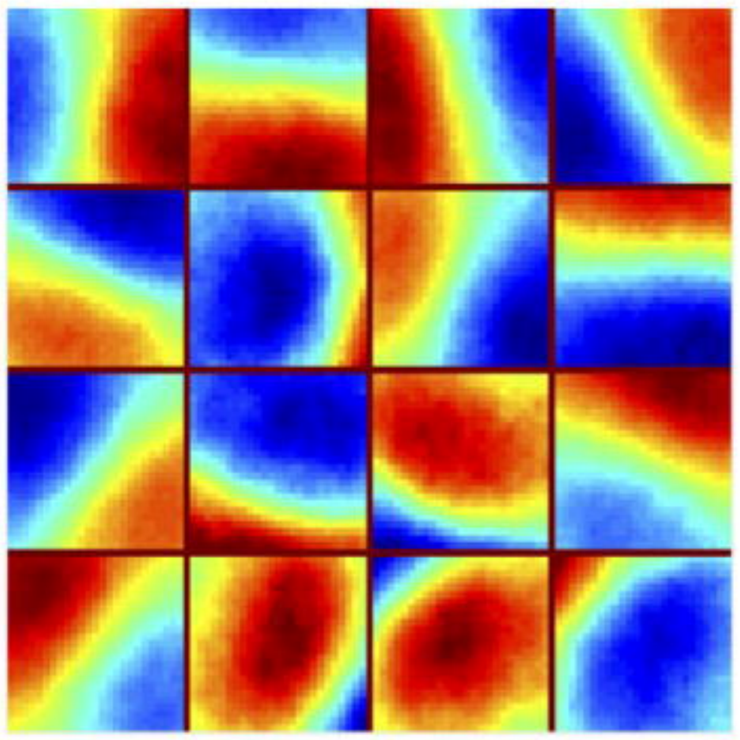


A

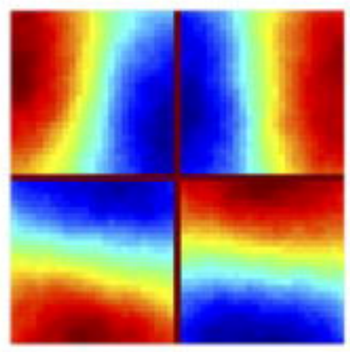
model with 36 lobula neurons



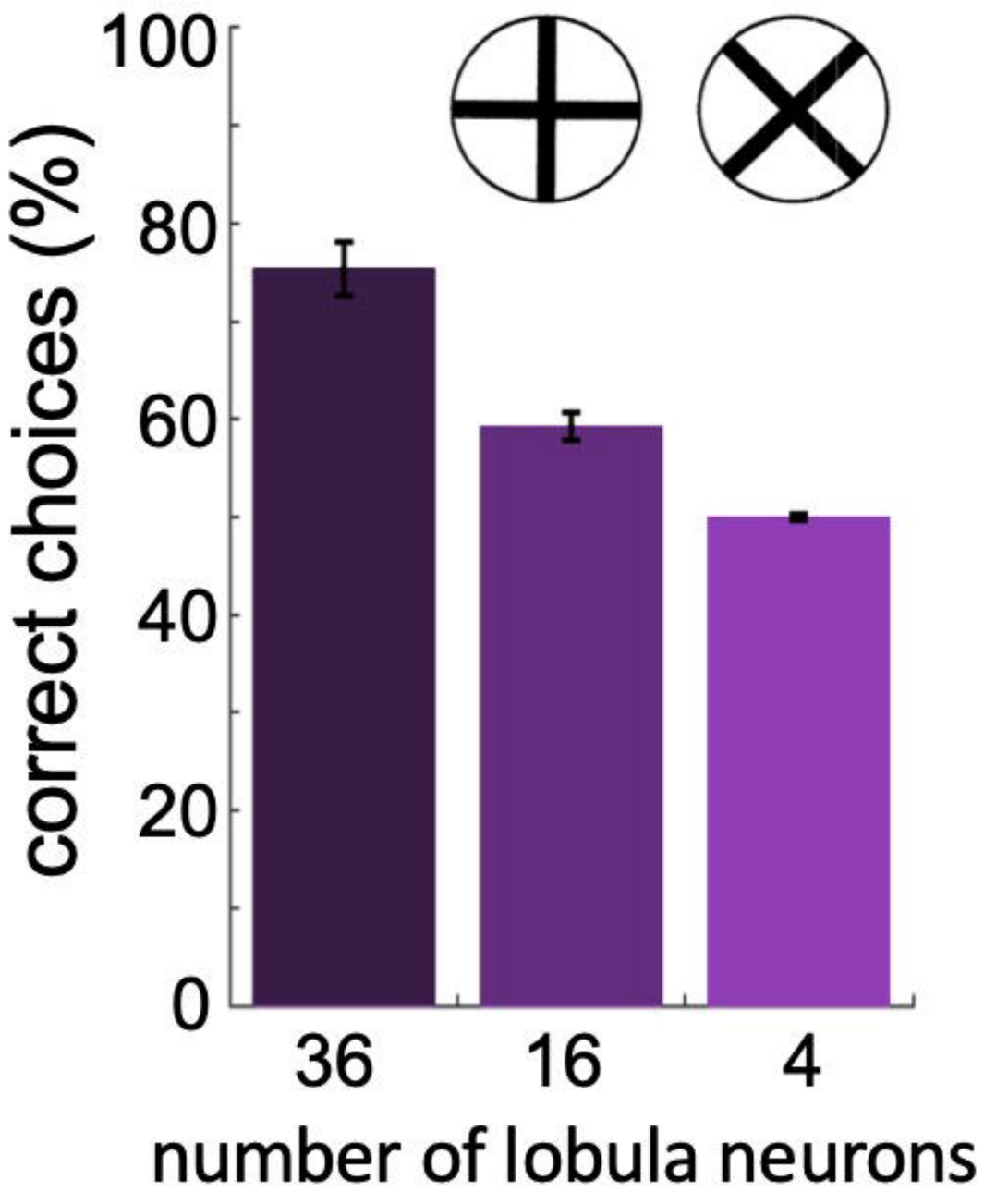
model with 16 lobula neurons



model with 4 lobula neurons



B



C