Abstract

One class of theoretical accounts of associative learning suggests that reinforcers are processed according to learning rules that minimize the predictive error between the expected strength of future reinforcement and its actual strength. The omission of reinforcement in a situation where it is expected leads to inhibitory learning of stimuli indicative for such a violation of the prediction. There are, however, results indicating that inhibitory learning can also be induced by other mechanisms. Here, we present data from olfactory reward conditioning in honeybees that show that (1) one- and multiple-trial backward conditioning results in conditioned inhibition (CI); (2) the inhibition is maximal for a 15-sec interval between US and CS; (3) there is a nonmonotonic dependency on the degree of CI from the US–CS interval during backward pairing; and (4) the prior association of context stimuli with reinforcement is not necessary for the development of CI. These results cannot be explained by models that only minimize a prediction error. Rather, they are consistent with models of associative learning that, in addition, assume that learning depends on the temporal overlap of a CS with two processes, a fast excitatory and a slow inhibitory one, both evoked by a reinforcer. The findings from this behavioral analysis of reinforcement processing are compared with the known properties of an individual, identified neuron involved in reinforcement processing in the bee brain, to further understand the mechanisms underlying predictive reward learning.

Introduction

One major concept in understanding the adaptive nature of associative learning is that animals acquire knowledge about predictive relations between either sensory cues or their own behavior and important reinforcing events. Predictive learning allows an animal to adapt its behavior to salient future events, such as rewards or aversive stimuli. In principle, two types of predictions can be learned: In excitatory learning, a cue or a behavioral act predicts the likely emergence of a reinforcer, thus resulting in the acquisition of positive predictions. Inhibitory learning results in the acquisition of negative predictions that anticipate the likely nonappearance of reinforcement. For example, in aversive learning, negative predictions might signal safety, or, in appetitive learning, foraging animals might learn to avoid nonrewarding food sources.

In classical conditioning procedures, animals learn predictive relations between events. Pairings of a discrete stimulus [conditioned stimulus (CS)] with an unconditioned reinforcing stimulus (US)
lead to the development of conditioned responses (CRs) to the CS that anticipate, and prepare the animal for, the following reinforcing event. In 1927 Pavlov showed that a CS can gain inhibitory properties [conditioned inhibition (CI)] resulting in the suppression of reinforcement related responses. CI can be demonstrated if acquisition of a CR during subsequent pairings of the putative inhibitory CS with the US is retarded (retardation of acquisition test) and if the inhibitory CS attenuates conditioned responding to another excitatory CS when presented in compound (summation test) (Rescorla 1967, 1969).

A classical experimental procedure that produces CI is the presentation of a stimulus in non-reinforced compound trials together with another stimulus that has been conditioned previously to the US. Here, learning to predict the nonoccurrence of a reinforcer appears to depend on the violation of a learned prediction. Inhibitory training does not, however, necessarily require two distinctive stimulus elements. For instance, backward pairing, where the CS follows the reinforcer, can also give rise to CI (Moscovitch and Loloerd 1968; Plotkin and Oakley 1975; Maier et al. 1976; Wagner and Larew 1985; Williams and Overmier 1990). Backward pairing does not usually produce excitatory conditioning (Rescorla 1988), but it has been shown that such excitatory conditioning occurs, especially after few backward pairing trials (Wagner and Terry 1976; for review, see Spetch et al. 1981). Several investigators demonstrated that backward pairing can result in inhibitory conditioning, with the interstimulus interval (ISI) being a critical factor for the degree of inhibitory conditioning to a CS (Plotkin and Oakley 1975; Maier et al. 1976; Wagner and Larew 1985). Whereas shorter ISIs result in inhibitory conditioning, longer ISIs are not effective. Thus, predictive learning appears to enable animals to extract the meaning of sensory events dependent on their temporal relationship to reinforcing stimuli.

How do animals establish predictions for future reinforcement and, in particular, what are the underlying processing principles that result in predictions for nonreinforcement? Most contemporary theories of classical conditioning (Rescorla and Wagner 1972; Mackintosh 1975; Pearce and Hall 1980; Holland 1990) suppose that an association between some internal representation of the CS and the US is formed, such that a CS builds up "associative strength" in evoking activity appropriate to the likely occurrence or nonoccurrence of the US. Theoretical accounts that make learning dependent on changes in reinforcement processing assume that learning reduces the predictive error for reinforcement. For example, in the Rescorla and Wagner (1972) model, the change in the associative strength of a stimulus in a certain trial depends on a reinforcement term

\[ r = \lambda - \Sigma V_i \]  

that minimizes the difference between the associative strength maximally supported by the actual US (\(\lambda\)) and the combined associative strength of all stimuli present at that trial (\(\Sigma V_i\)). Thus, \(\lambda - \Sigma V_i\) computes a prediction error for reinforcement. According to this rule, inhibitory training requires that a stimulus is presented without reinforcement (\(\lambda = 0\)) in the presence of other stimuli with positive associative strength. The stimulus will then gain negative associative strength, because the reinforcement term \(r = -\Sigma V_i\) is negative. Stimuli presented in a situation without a US become conditioned inhibitors if either discrete stimuli, as in the classical procedure for CI, or background information (e.g., contextual cues of the experimental situation) predict the occurrence of a US.

Although the Rescorla and Wagner (1972) model explains several conditioning phenomena, as a trial-based model, which updates the associative strength during a trial independent of the temporal arrangement of the stimuli, it fails to explain the dependence of conditioning on the ISI between CS and US. This limitation led to the development of real-time models that account for the ISI dependence of associative learning (for an overview, see Malaka and Hammer 1996). In particular, Sutton and Barto’s (1990) time-derivative (TD) model is able to predict the ISI dependence of both excitatory and inhibitory learning (Malaka and Hammer 1996). Moreover, the TD model has been used recently in several variants to describe the properties of neural systems that process reinforcement and to propose putative neural implementations of predictive learning (Barto 1995; Houk et al. 1995; Montague et al. 1995, 1996; Moore and Choi 1997; Schultz et al. 1997; for a similar approach on value processing, see also Friston et al. 1994). In the tradition of the Rescorla and Wagner (1972) model, the TD algorithm makes inhibitory learning dependent on the presence of context stimuli with positive associative strength (Malaka and Hammer 1996).
The TD reinforcement term

\[ r(t) = \lambda(t) + \gamma Y(t) - Y(t-1) \]  

combines a term that is driven by the US \( \lambda(t) \) with the first time derivative of the learned prediction of reinforcement \( Y(t) \). This time derivative is computed by taking the difference between the amount of reinforcement predicted at a given time step \( \gamma Y(t) \) minus that predicted at the preceding time step \( Y(t-1) \). The factor \( \gamma \) \( (0 \leq \gamma \leq 1) \) reflects the uncertainty of \( Y(t) \), because this actual prediction of reinforcement can only be estimated by using information provided by the associative strength taken from \( t - 1 \) for stimuli present at the actual time step \( t \) (Sutton and Barto 1990). The TD model explains the ISI dependency of backward inhibitory learning only if context information is eligible for entering into associative learning during US presentation (Malaka and Hammer 1996). After the US, the ongoing presence of the experimental context results in a positive prediction of reinforcement and in a negative reinforcement term \( \gamma Y(t) - Y(t - 1) \). Negative reinforcement, in turn, extinguishes the associative strength of the context and produces negative associative strength for those stimuli presented after the US. Because the degree of associative strength of the context and, hence, that of negative reinforcement decreases with time, the TD model predicts a degree of CI that is maximal directly after US offset and monotonically declines with longer ISIs.

In contrast, another theoretical account for the ISI dependence of associative learning, Wagner's (1981) SOP model, does not require the association of context information for backward inhibitory learning. Rather, it postulates two internal processes driven by reinforcing events. In the tradition of earlier theories (Konorski 1948; Solomon and Corbit 1974; Schull 1979), SOP assumes that a US evokes two antagonistic modes of activity, an early, fast decreasing, excitatory one and a later, longer lasting, inhibitory one (Wagner 1981; Wagner and Larew 1985), which combine to a biphasic internal reinforcement signal. Dependent on the overlap of a CS with this reinforcement signal, net excitatory or inhibitory learning will occur. In particular, SOP predicts a nonmonotonic ISI dependence of backward inhibitory learning: Long and very short ISIs result in no inhibitory learning; inhibitory learning is maximal with short ISIs. The two SOP processes can be regarded as dividing learning about events that occur during epochs of reinforcement into two periods: a brief one that precedes or is concurrent with reinforcement and one that follows reinforcement. Events that occur during the first period will gain a positive predictive value; those that follow, a negative one.

In summary, the different theoretical accounts suggest that the capacity of an animal to learn both positive and negative predictions depends on the means by which brains process important reinforcing events. Behavioral analysis can, however, serve to unravel the constraints of reinforcement processing that enable predictive learning. In honeybees, the proboscis extension response (PER) can be conditioned to an odor or mechanical cue (CS) by even a single pairing of the odor with sucrose solution applied to the antennae and/or proboscis (US) (Bitterman et al. 1983; Menzel 1990). This appetitive learning depends strongly on the CS/US ISI (Menzel and Bitterman 1983). Unpaired training, in which CS and US are separated by long ISIs, leads to a subsequently retarded acquisition of the CR (probability of PER) during forward conditioning (Bitterman et al. 1983), which depends on the number of unpaired trials (Menzel 1990). This indicates that unpaired training may lead to inhibitory conditioning of the CS. Preliminary data also suggested that one-trial backward pairing of US and CS results in inhibitory learning of the CS (Hellestem and Hammer 1994).

The neural basis of the reinforcing function of the sucrose reward has been resolved to some extent (Hammer 1993, 1997; Hammer and Menzel 1995). A single identified neuron, the VUMmx1 neuron, responds to the reward, and its activity is sufficient to substitute the reward in a single forward pairing of an odor and activity of VUMmx1. Backward pairing, in which the odor is applied after onset of VUMmx1 activity, is not effective. After differential conditioning, in which one odor has been rewarded, whereas another odor has been presented unrewarded, the rewarded odor, but not the unrewarded, evokes a long-lasting excitation of VUMmx1 (Hammer 1993). This experience-dependent plasticity of the response properties of VUMmx1 makes it a good candidate for the investigation of how and whether theoretical models, which make learning dependent on changes in reinforcement processing, are implemented in nervous systems (Montague et al. 1995; Hammer 1997). The finding that reward-predicting odors fire VUMmx1 is, however, not compatible with the notion that VUMmx1 implements the reinforcement term of the Rescorla and Wagner model.
Hellstern et al. (Hammer 1997). This notion would predict that reward-predicting odors inhibit VUMmx1, because in the absence of a reinforcer (λ) the reinforcement term of the Rescorla and Wagner model, \( r = \lambda - \sum V_i \), becomes negative. Moreover, it makes it difficult to understand how bees should show CI in paradigms in which a neutral stimulus is paired with another stimulus that has positive associative strength. In that case, the neutral odor would be paired with increased activity of VUMmx1, which should result in excitatory learning (second-order conditioning).

Therefore, we investigated the basis of inhibitory learning in PER conditioning using a behavioral approach. Because backward inhibitory learning should allow us to distinguish between two principal modes of reinforcement processing, we examined whether backward pairing of an odor and the sucrose reward with different ISIs gives rise to inhibitory learning of this odor. In a further experiment, we analyzed the influence of the experimental context on CI induced with a single backward pairing of CS and US.

Our results show that backward training of CS and US leads to CI with a nonmonotonic dependence on the ISI. Moreover, comparison of the experimental data with the different theoretical accounts suggests that backward inhibitory conditioning in bees requires a second inhibitory US process. On the basis of these findings, we shall discuss putative neural implementations of reward processing in the brain of the honeybee.

**Experiment 1**

**Introduction**

In a first experimental series, we examined whether backward pairing of CS and US with different ISIs results in inhibitory learning in PER conditioning. A demonstration of CI requires two paradigmatic tests: the retardation of acquisition test and the summation test (Rescorla 1967). As a retardation test, we used a single backward pairing of CS and US with different ISIs followed by a single forward pairing to investigate the influence of backward pairing on the acquisition of a CR. As a summation test, we investigated the effect of an inhibitory CS established during backward pairing on the strength of the CR in a compound trial with another, excitatory, CS. In contrast to the retardation experiment, a multiple-trial procedure was used to rule out the possibility that backward inhibitory learning is only a single-trial effect. To allow for a graded measure of the PER, electromyogram recordings were taken from muscle M17, which monitors the whole PER and was used for measuring experience-dependent changes of the motor program underlying PER (Rehder 1987; Smith and Menzel 1989; Braun and Bicker 1992; Hammer et al. 1994).

**Materials and Methods**

**PREPARATION OF ANIMALS**

Worker bees (*Apis mellifera*) were caught at the hive entrance the afternoon before the experiment, immobilized by cooling, and subsequently fixed in metal tubes with strips of tape placed between head and thorax. In addition, the heads were fixed at the tape with a drop of molten wax (Eicosan) allowing free movement of the antennae and the mouthparts. After recovering from the cooling, the bees were fed with a 30% sucrose solution and stored overnight at room temperature in a dark, humid place. On the day of the experiment, a metal wire (50 μm in diam., material V2A steel, and insulated, except for the tip) was inserted into the head capsule close to the insertion point of muscle M17. An indifferent electrode was inserted into the head capsule close to the insertion point of muscle M17. An indifferent electrode was inserted into the contralateral capsule close to the insertion point of muscle M17. An indifferent electrode was inserted into the contralateral capsule close to the insertion point of muscle M17. An indifferent electrode was inserted into the contralateral capsule close to the insertion point of muscle M17. An indifferent electrode was inserted into the contralateral capsule close to the insertion point of muscle M17.

**STIMULATION**

Odor stimuli (CS) were applied by an aquarium pump, which blew air through a syringe that contained a piece of filter paper soaked with an odorant (3 μl). A plastic tube connected to the
REINFORCEMENT PROCESSING IN HONEYBEES

syringe directed the air stream to the bee’s head (5-cm distance). The duration of the odor stimuli (6 sec) was controlled by magnetic valves opened and closed by a timer. The accumulation of scented air in the experimental chamber was prevented by the exhaust. The sucrose reward (US) was applied by first touching the antennae and then the proboscis with a wooden toothpick moistened with 30% sucrose solution. The duration (2 sec) and time of sucrose stimulation was guided by an electronically controlled auditory signal. A trial lasted for the time bees spent in the experimental position before, during, and after stimulation. For a standard trial bees were brought into the experimental position 25 sec before odor or sucrose presentations and were kept in this position for at least 21 sec after stimulation so that the overall duration of a trial was 53 sec.

PROCEDURES

RETARDATION OF ACQUISITION

In this experiment we investigated the effect of the ISI during a single backward pairing of CS (carnation) and US on a subsequent forward pairing of CS and US (Fig. 1A). In a first phase of the experiment, all bees received a single presentation of the US to control for the quality of muscle recordings. Bees that showed no PER or had a poor electromyogram were discarded. The experiment started with an odor stimulus (pretest), which was presented 30 min after the sucrose control and 30 min before training. Training consisted of either a single trial with a US presentation (USonly group) or a presentation of the US followed by a CS at different time intervals (6, 15, 30, 60, 120, 300, and 600 sec; backward pairing groups). This resulted in different trial durations for the various interval groups. During backward pairing, only bees that received US and CS at an interval of 300 or 600 sec received either stimulus in a separate trial (see text). Bees with a US–CS ISI of 300 or 600 sec received either stimulus in a separate trial (broken line in exp. pos.). The duration a bee remained in the exp. cont. constituted a trial (see text). Bees with a US–CS ISI of 300 or 600 sec received either stimulus in a separate trial (broken line in exp. pos.). (B) Summation test after three forward pairings of odor A with an ITI of 1 min followed (8 min later) by three backward pairing trials with different ISIs (6-min ITI) with odorant B. The ISI during a Backward pairing resulted in a variable interval between the CS and the US of the next trial. The effect of training was determined by a randomized presentation (2-min ITI) of each of the single odors A and B and the compound AB, which started 32 min after the last Backward pairing (summation test). The individual bees remained in the exp. pos. throughout the entire training.

SUMMATION TEST

In this experiment we investigated how pre-training of one odorant with different CS-US backward intervals affected the responsiveness of bees to another odorant that had been conditioned previously with CS-US forward intervals, when both odorants were presented in a compound test trial after training (Fig. 1B). During training, each bee was conditioned with three forward pairings of odorant A with an ISI of 5 sec and an inter-trial interval (ITI) of 1 min. Eight minutes later, three backward pairing trials (ITI of 6 min) with odorant
B followed. During backward training, different groups of bees received US and CS with an ISI of either 6, 15, or 300 sec. Bees were placed in the experimental position 25 sec before the first CS presentation of the forward pairings and stayed there until 25 sec after the last CS of the backward pairing phase. During tests, which started 32 min after the last backward pairing, each bee received a randomized presentation of the individual odorants A, B, and their binary mixture AB with an ITI of 2 min. The two odorants, geraniol and 1-hexanol, were used in a balanced manner.

DATA ANALYSIS AND STATISTICS

Recorded electromyograms were stored on tape or via an A/D converter (CED1401, “Spike2 data capture” program; Cambridge Instruments) directly onto a computer hard disk. The number of recorded M17 spikes per stimulus was determined during a 10-sec interval after stimulus onset (“Spike2 analysis” program). Differences between the number of M17 spikes in the various experimental groups were determined by the Kruskal-Wallis H test, whereas between-group differences were determined by the Mann-Whitney U tests. The Wilcoxon test was used to examine within-group differences in the summation test.

Results

RETARDATION OF ACQUISITION

In this experiment we investigated whether a single backward pairing of CS and US attenuates the acquisition of a CR to that CS in a subsequent forward pairing trial. If so, the response to the CS should be reduced compared with a control.

Of 369 bees that passed the initial test for a good recording quality, 153 either died or the recording quality decreased throughout the experiment, ruling out an evaluation of electromyograms. These bees were discarded from further analysis. Of the remaining 216, several responded weakly to the odor during the pretest, as revealed by the odor-evoked activity of muscle M17. The number of M17 spikes, however, did not differ between experimental groups [median, between 0 spikes and 16 spikes; \( t = 5.80, df = 7 \), not significant (NS), \( H \) test], indicating no group-dependent difference in the odor-evoked response before training.

Whenever bees were moved into the experimental position, only four proboscis extensions occurred before stimulation with either the US or CS (place-ment response). After training, there was a clear increase in the odor-evoked number of M17 spikes (median, between 16 and 211). A comparison of the median of M17 spikes for the different groups showed that the strength of the response of the bees to the odor during the test depended on their experimental history (Fig. 2) \( (t = 14.93, df = 7, P < 0.05) \). Moreover, the number of M17 spikes evoked by the test odor showed a nonmonotonic dependence on the US-CS ISI during the backward trial. It decreased from bees that received a backward training with a very short ISI (6 sec) to those that were trained with a short interval (15 sec) and increased again with longer ISIs. Group-by-group comparison revealed that the median of M17 spikes, evoked by the test odor, of those bees that
received a backward pairing of US and CS with an ISI of 15 sec was significantly lower than that of bees that were trained with an ISI of 300, 600, or 6 sec (see Fig. 2; \( P < 0.01 \) for 15 sec against 600 sec, all others \( P < 0.05 \), U test). Moreover, the median of M17 spikes evoked by the test odor for bees that were trained in the backward trial with an ISI of 15 sec, but not for bees that were trained with ISIs of 6, 300, or 600 sec, was significantly lower than that of the “USonly group” \( (P < 0.01) \). This indicates that pretraining with a backward pairing of US and CS inhibits the subsequent acquisition of a CR to this CS in a forward pairing only for ISIs in a certain range.

SUMMATION TEST

In the second experiment we investigated whether pretraining with backward pairings of CS and US with different ISIs not only retards acquisition of a CR to that CS but also attenuates conditioned responding to another, excitatory, CS. If so, the response to the excitatory CS should be reduced when presented in compound with the putative inhibitory CS.

Throughout this experiment, 79 of 148 bees either did not survive the experiment or were discarded because of poor recording quality. During the entire experiment, only 3 bees showed a placement response, and, during the test, as revealed by the electromyograms, 57 bees responded to the odor that was trained with forward pairings with the US (odor A). To the odor that was trained with backward pairings (odor B), M17 activity was recorded in only 12 bees during the test, and the number of M17 spikes was independent of the ISI during backward pairings of odor B \((t = 0.134, df = 2, \text{NS}, H \text{ test})\). As shown in Figure 3A, the median of M17 spikes evoked by the test presentation of odor A also did not depend on the ISI (6, 15, or 300 sec) with which bees were trained during backward pairings of the US and odor B (NS). Thus, backward pairings of one odor with different ISIs did not affect the ability to show a CR to the other odor. These results were independent of which odor was chosen as A or B. When both odors were presented in compound (AB), the ISI during backward pairings had a significant effect on conditioned responding (Fig. 3B). The median of M17 spikes evoked by presentation of the compound AB for the group of bees that received backward pairings of US and B with an ISI of 15 sec was lower than that for the group with an ISI of 6 or 300 sec \((t = 6.01, df = 2, P < 0.05)\). This indicates that backward pairing with a 15-sec ISI suppresses the response to the compound AB as compared with the 6- and 300-sec ISIs.

**Figure 3:** In a summation test after multiple backward pairings of CS and US, the response to a compound of this CS (odor B) and an excitatory CS (odor A) is suppressed dependent on the backward ISI. (A) The responses to the excitatory odor A did not differ between groups that were trained with different backward ISIs for odor B. The numbers below abscissa indicate the ISI (in sec) used during backward pairings. Shown are the numbers of odor-evoked M17 spikes (median with interquartile ranges) in an extinction trial with odor A. (B) The responses to the compound of odor A and B is smallest for the intermediate ISI of 15 sec \((df = 2, t = 6.01, P < 0.05, H \text{ test})\). Shown are the numbers of odor-evoked M17 spikes (median with interquartile ranges) in an extinction trial with the compound of odor A and odor B. Groups are named as in A. (C) The presentation of the compound of odor A and B reduces the response for the intermediate ISI and augments the response for shorter and longer ISIs as compared with the response to odor A alone \((df = 2, t = 15.11, P < 0.005, H \text{ test})\). Shown are the calculated differences between the numbers of odor-evoked M17 spikes following a presentation of the compound (AB) minus the number of spikes evoked by a presentation of odor A alone (median with interquartile ranges). Groups are named as in A. The numbers of bees are 28 for group 6, 21 for group 15, and 20 for group 300.
To determine whether this effect was caused by an attenuation of the response to A in the presence of B, we calculated the difference between the number of M17 spikes evoked by AB minus those that were evoked by presentations of A alone (Fig. 3C). For none of the three groups is this difference significant (6-sec group: $z = 1.90$, $P = 0.057$; 15-sec group: $z = 1.69$, $P = 0.092$; 300-sec group: $z = 1.89$, $P = 0.057$; Wilcoxon test). However, the presence of B, in fact, reduces the number of M17 spikes evoked by the AB compound as compared with A alone in the case of backward training of B with an ISI of 15 sec but increases the number of M17 spikes for backward training of B with ISIs of 6 or 300 sec. This antagonistic effect yielded a highly significant between-group difference ($t = 15.11$, $df = 2$, $P < 0.005$, $H$ test). Thus, backward pairing with a 15-sec ISI appears to be sufficient to suppress response to the AB compound when compared with groups trained with a 6- or 300-sec backward ISI that showed an overall increase in response to the compound.

Discussion

The finding that backward pairing of CS and US, depending on the ISI, attenuates the acquisition of a CR in a subsequent forward pairing trial of CS and US (retardation test) indicates inhibitory backward learning in bees. However, adding an odor B that was trained with a backward ISI of 15 sec, which was most effective in the retardation test, fails to significantly reduce the response level to a conditioned excitor (A) in an AB compound, a result that would normally be indicative of the demonstration of CI in a summation test. Rather, in our experiment, the presumed inhibitory effect of training with the 15-sec backward ISI in the summation test is evident when the response level to the AB compound is compared with that of groups that received a backward training with shorter (6-sec) and longer (300-sec) ISIs. Thus, backward pairing of a US and CS results in both effects, which are critical for the demonstration of CI (Rescorla 1967).

Backward training with the 15-sec ISI inverted the tendency of an increased response strength to the AB compound, which occurred after backward training of B with the 6- and 300-sec ISI (Fig. 3C). This tendency may be explained by the higher total odor concentration in the compound trial. Alternatively, odorant B might have gained positive associative strength during training either directly or indirectly from generalization from odorant A to B, or both. The long forward ISI (60 sec) between the presentation of B and the next following US makes excitatory conditioning of B unlikely in the case of backward pairing with an ISI of 300 sec. Excitatory conditioning might have occurred during backward pairings with an ISI of 6 sec. Such excitatory conditioning for short backward ISIs was found in eyelid conditioning in rabbits (Wagner and Larew 1985). The lack of evidence for excitatory learning with the 6 sec ISI in the retardation experiment, however, suggests that generalization may have caused a transfer of excitatory learning from A to B. This notion is consistent with the high level of initial generalization normally observed in PER conditioning in differential conditioning procedures (Bitterman et al. 1983; Menzel 1990; Hammer and Menzel 1995). Thus, B would gain positive associative strength caused by pretraining with A that, in turn, may interfere with inhibitory learning of B during backward training. A transfer of excitatory learning to odorant B could have reduced the net degree of inhibitory learning in all ISI groups and may have led to a net positive associative strength of B for the 6- and 300-sec ISI after training. Under this assumption, a response threshold and a non-linear translation of associative strength into response strength would be required to account for the observed effects, because, for all backward ISIs, only a few bees responded to B in the test trials.

Experiment 2

Introduction

The nonmonotonic ISI dependence of backward inhibitory learning in bees could, in principle, be explained by two alternative real-time models of associative learning, one of which, the TD model (Sutton and Barto 1990), requires the association of context information (Malaka and Hammer 1996), whereas the other, the SOP model (Wagner 1981), postulates two internal processes driven by reinforcing events.

In our experimental design, two sources of context information might have acquired positive associative strength during US presentations: First, stimuli of the experimental chamber, in which bees were kept throughout the experiment, such...
as visual and mechanical cues, luminescence, smell, etc., and second, stimuli associated with the position in which bees received stimulation, such as the air stream provided by the exhaust. For several reasons, however, the contribution of context information might have been weak. In the retardation experiment, stimuli provided by the chamber did not consistently predict reinforcement, because all the bees of one experimental session were kept in and moved within the chamber throughout the entire period of training and testing. This should have prevented learning of this source of context information, although it does not rule it out. In contrast, stimuli of the experimental position did consistently predict reinforcement, because bees were moved into this position 25 sec before US delivery in the retardation experiment. Thus, bees might have associated these stimuli with reinforcement. Bitterman et al. (1983) reported that bees develop proboscis extension responses when placed in front of an exhaust during an unpaired training with an olfactory CS and a sucrose reward. Consistently, they suggested that learning of context information might account for the retardation of acquisition after multiple-trial unpaired training. However, in their experiment, bees were, in contrast to our experiments, placed in front of the exhaust only 6 sec before the US. This short interval might have facilitated learning of the experimental context. Our finding that only a small number of bees, if any, responded independently of CS or US delivery suggests that, in our experiment, learning about the predictive value of positional cues was weak. Moreover, in the summation experiment, we reduced the predictive value of this source of context information, because individual bees remained in the experimental position throughout the whole training procedure.

An additional observation of the retardation experiment argues against a strong contribution of context information. In two of the experimental groups, the groups with either a 300- or a 600-sec CS-US backward ISI, bees stayed in the experimental position in front of the exhaust between US and CS for only 45 sec. The TD model predicts that the degree of CI depends on the associative strength of the context and, hence, the amount of negative reinforcement evoked by this context at the time of CS delivery. Because the ongoing presence of the context would result in extinction of associative strength, bees in these two groups should have shown a level of CI comparable with those in groups with shorter ISIs (e.g., the 30-sec and 60-sec group). This, however, was not the case.

Because these observations do not rule out that other sources of context information, for instance those associated with the experimental chamber, may contribute to backward inhibitory learning, we designed a further experiment that explicitly removed all of the context stimuli for a certain period between US and CS for training with a backward ISI of 300 sec. In comparison with bees that experienced the experimental context throughout the entire US-CS ISI, these bees should show considerable inhibitory learning, if learning of the context is the dominant source for CI. Moreover, to maximize a putative contribution of the context, bees in the group that experienced the change of the context stayed in the experimental position between US and CS for only 15 sec, the ISI that was most effective in causing inhibitory learning.

Materials and Methods

Bees were caught from an indoor flight facility and were otherwise treated and prepared as in the first experiment. They were mounted for stimulation in the experimental chamber, but no muscle recordings were taken. As a second, control context, a small dark box (20 x 15 x 10 cm) was used, which was mounted in the corridor outside the laboratory. The box was always closed when a bee was placed inside. Duration and application of CS and US were the same as in the first experiments.

PROCEDURE AND DATA ANALYSIS

The experimental design used was the same as for the retardation of acquisition test. After a pre-test trial with the CS, one backward pairing trial of US and CS was followed by one forward pairing trial and a subsequent test trial. The ITI was 30 min. The four experimental groups differed only in their handling during the backward pairing phase. One group received a trial with a single US presentation (USonly). One group received a backward pairing trial of US and CS with a interval of 15 sec and stayed in the experimental position in front of the exhaust throughout (group 15/15). For the other two groups, the US-CS interval was 300 sec. One group (300/300) remained in the experimental position for the whole 300 sec, whereas the
other (group 300/15) remained in front of the exhaust for only 15 sec and was placed into the control context (dark box) for the remaining 285 sec (see Fig. 4).

In contrast to experiment 1a and 1b, the effects of training were determined by calculating the proportion of bees that showed an odor-evoked PER from the total of bees tested. This procedure seemed reasonable because the results from the first experiments indicated that the graded measures of the response strengths through muscle recordings were not necessary to detect between-group differences. A response was scored as PER when the tip of the proboscis crossed a virtual line between the opened mandibles. Between-group differences were determined using the Freeman test (Freeman and Halton 1951) and the $\chi^2$-tests.

Results

Those 190 out of 335 bees that did not die and showed an unconditioned response (UR) after the test trial at the end of the experiment were analyzed. Only two of these bees showed a placement response. The different handling of the four groups resulted in an overall significant effect on the response level (%PER) to the test odor ($f_{max} = 749.44$, $P < 0.002$, Freeman test). As shown in Figure 5, backward pairing with an ISI of 15 sec (group 15/15, $n = 49$) produced an inhibitory effect. The test response level of group 15/15 was significantly lower than that of group USonly ($n = 52$, $\chi^2 = 15.05$, $P = 0.0001$) and than that of both groups that received backward training with an ISI of 300 sec (group 300/300: $n = 45$, $\chi^2 = 4.09$, $P < 0.05$; group 300/15: $n = 44$, $\chi^2 = 4.5$, $P < 0.05$). The response level of these two groups did not significantly differ from that of group USonly (300/300 vs. USonly: $\chi^2 = 3.32$, NS; 300/15 vs. USonly: $\chi^2 = 2.89$, NS). In contrast to the prediction of the TD model, the response level between the group that remained in the experimental position (group 300/300, %PER = 51%) and the group that experienced a change in the context (group 300/15, %PER = 52%) did not differ. This shows that the critical factor controlling the degree of CI is the interval between US and CS in the backward pairing phase and not the time of context exposure.

Discussion

These results suggest that, at least in the single-trial backward pairing paradigm of our experiment, learning of the experimental context is not responsible for CI. The three experimental groups
(USonly, 15/15, and 300/300), which were comparable to the respective groups in the first retardation experiment, showed a response pattern similar to that of the first experiment. Fewer animals in group 15/15 responded to the test odor than in the USonly group indicating that one-trial conditioning was retarded. Moreover, the response level of the critical group 300/15 was the same as that for group 300/300, which received the same US-CS interval, and was higher than that for group 15/15, which spent the same time in the experimental context. If backward inhibitory learning of a stimulus depends on its occurrence in a context that has acquired positive associative strength, the time that bees have spent in this context between the presentation of US and CS should be the only critical factor for the degree of inhibitory conditioning. That is, the response level for group 300/15 should have approached that of group 15/15. Our finding to the contrary suggests that context stimuli in backward inhibitory learning can not, at least in our experimental setup, play a significant role.

**General Discussion**

Our experiments demonstrate that single- and multiple-trial backward pairings of CS and US are sufficient to induce CI in olfactory PER conditioning in honeybees. This backward inhibitory learning varies nonmonotonically with the US-CS ISI. It was neither detectable for very short (6 sec) nor for longer ISIs (>120 sec), whereas the inhibitory effect was maximal with short ISIs (15 sec). Moreover, the results of experiment 2 indicate that there is no measurable effect of stimuli associated with the experimental setup on the degree of CI, suggesting that the temporal relationship between US and CS is sufficient to produce and control the degree of inhibitory learning.

The finding that CI decreases with longer ISIs corresponds to the monotonic ISI dependence of CI in a shuttle avoidance paradigm (Maier et al. 1976). However, these experiments showed no indication of a reduction of CI with very short ISIs. In contrast, a nonmonotonic dependency of CI has been reported for eyelid conditioning in rabbits with a ISI of 31 sec compared with a 1- or 600-sec ISI (Wagner and Larew 1985). Because a successful demonstration of a nonmonotonic dependence of backward CI might depend on the appropriate choice of short ISIs and of the duration of the stimuli, it is, at present, an open question whether this is a general feature of conditioning. The fact that backward inhibitory learning with a nonmonotonic ISI dependence occurs in paradigms as different as aversive eyelid conditioning in rabbits and appetitive olfactory PER conditioning in honeybees indicates, however, that this might be the case.

Inhibitory learning requires that a stimulus overlaps with a mechanism that results in negative internal reinforcement. Our finding that CI decreases with the time elapsed after presentation of a reinforcer suggests that this mechanism is triggered by the reinforcing event itself and decreases over time. Therefore, trial-based theoretical mechanisms, like those proposed by Rescorla and Wagner (1972), Mackintosh (1975), or Pearce and Hall (1980), are not sufficient. Moreover, for several reasons, real-time accounts like the TD algorithm (Sutton and Barto 1990), which produce a negative internal reinforcement through the ongoing prediction of reinforcement provided by the experimental context, are also not sufficient to explain backward inhibitory learning in bees. The reinforcement term of the TD model (see equation 2) predicts maximal inhibitory learning for stimuli that occur directly after US offset, because then the US-driven term λ(t) equals zero, and the positive associative strength of context stimuli produce a strong negative internal reinforcement that decreases with time. In our experiments, however, no CI could be detected for the shortest backward ISI of 6 sec. This might suggest that the US-driven positive internal reinforcement λ(t) should outlast and slowly decay after stimulation. For stimuli presented at the time when this process equals the amount of negative reinforcement produced by the associative strength of context stimuli, no inhibitory learning would occur, whereas later presentation would result in inhibitory learning. Even though this slight modification of the TD model would predict a nonmonotonic ISI dependence, two other observations argue against a major contribution by the association of context stimuli and, hence, the mechanism of the TD model for backward inhibitory learning. For a large inhibitory effect after only a single backward pairing of CS and US, the TD model would require a large degree of context conditioning. This, in turn, would predict some response to the experimental context, a prediction not confirmed by our experimental results. Our finding in experiment 2 that a change in the context after US delivery, which should prevent its...
Extinction, did not influence the degree of CI, suggests that learning about the experimental context can be neglected.

We, therefore, propose that the generation of both positive and negative internal reinforcing signals is an inherent property of the processing of important events, independent of prior associative experience. This suggests that backward conditioning of CS and US should be a standard procedure for inhibitory conditioning and that the learning of positive and negative predictions is a basic feature of the association process. This view, moreover, consistent with those theoretical accounts that, in contrast to the TD model, assume that conditioning depends on the overlap of a CS with two US-evoked processes, an excitatory and an inhibitory one (Konorski 1948; Schull 1979; Wagner 1981). Joint activity of a CS with a second, inhibitory, US process (fall of activity, b-state, A2-process) leads to inhibitory conditioning in these models. The net degree of excitatory and inhibitory learning depends entirely on the temporal relationship between CS-evoked activity and the two US processes. The results of our experiments, mainly the retardation experiment, might enable us to estimate the time courses of these processes and, hence, of the internal reinforcing signal.

First, maximal inhibitory effects with an ISI of 15 sec require a maximal inhibitory process that peaks at this time after US delivery. Second, the nonmonotonic ISI dependence of inhibitory learning can be explained if a certain ratio for the strength of the inhibitory to the excitatory US process is assumed for the 6-sec ISI. Either both are weak or excitatory conditioning supported by an excitatory reinforcement signal must compensate for inhibitory conditioning that results from the increasing strength of the inhibitory US process. Third, the inhibitory process should express a slow, long-lasting, falling phase that decays within minutes.

Although context associations are not required for backward inhibitory learning in bees, our results do not exclude the possibility that context learning could facilitate inhibitory learning. Similarly, even though our experiments cannot be described by error reduction learning rules, rules that compute a predictive error for future reinforcement may be required for the explanation of other important features of olfactory PER conditioning. Several investigators have proposed that both the learning of floral parameters as signals for the profitability of food sources (Couvillon and Bitterman 1985, 1986, 1988; Greggers and Menzel 1993; Montague et al. 1995) and olfactory PER conditioning (Hammer 1997; Smith 1997) are consistent with the basic assumptions of error reduction learning rules. For example, in PER conditioning, unpaired presentation of an olfactory CS and the sucrose US with ISIs of 5 min retards the acquisition of that CS (Bitterman et al. 1983) where the strength of the retardation effect depends on the number of unpaired trials (Menzel 1990). In this paradigm, evidence for the association of contextual stimuli (see Bitterman et al. 1983) indicates that the processing of a predictive error might account for inhibitory learning. If CI depends entirely on an inhibitory process evoked by the US, this process would have to last for at least 5 min to produce inhibitory learning in the unpaired paradigm. This seems, however, to be unlikely, because we did not find significant inhibition for this ISI. Our experiments (in particular the summation test) were optimized to reduce the possibility of putative context conditioning, whereas the protocol used by Bitterman et al. (1983) may have favored learning about the context, because bees where placed into the training position 6 sec before stimulation. Therefore, both experimental designs could have favored different mechanisms for inhibitory learning. An inhibitory internal reinforcement evoked by US presentation only and negative reinforcement produced by stimuli (including contextual cues) that have acquired associative strength are, however, not mutually exclusive. For instance, the SOP model assumes that conditioning results in the ability of a CS (or contextual cues) to evoke the inhibitory US-related process. This feature of the model accounts for inhibitory learning that is driven by the presentation of a stimulus in a situation in which other stimuli predict reinforcement. Because SOP postulates that activation of the inhibitory process reduces the ability of the US to evoke the excitatory process (Wagner 1981), this feature of the model is equivalent to the reduction of a predictive error. In summary, behavioral analysis of olfactory PER conditioning suggests that bees are able to make predictions about rewards based on two factors: internal features of reward processing independent of prior associative learning and mechanisms that result in predictive error reduction. Although the SOP model combines these two features, our results do not necessarily suggest that it is directly applicable to PER conditioning and is, therefore, implemented within the brain of the honeybee.
One major approach for understanding the neural basis of behavioral learning, which has recently been applied to several learning paradigms, is to investigate how real-time computational models might be implemented by neural circuits that are involved in learning. For both the SOP model and the TD model, implementations within the cerebellar circuit that support motor learning have been suggested [SOP (Wagner and Donegan 1989); TD (Moore and Choi 1997)]. The TD model, moreover, directly predicts features of the mammalian reward processing system (Barto 1995; Houk et al. 1995; Montague et al. 1996; Schultz et al. 1997), and a simplified version of TD has been suggested to predict properties of the reward processing VUMmx1 neuron in the honeybee, as well as predictive learning during foraging (Montague et al. 1995; see also Hammer 1997). One property of the VUMmx1 neuron of special interest in this context is that not only sucrose rewards activate VUMmx1 but also odors that have been learned to predict sucrose (Hammer 1993). As mentioned above, this rules out that VUMmx1 implements the reinforcement term (equation 1) of the Rescorla and Wagner model (1972). It is also not compatible with a basic assumption of the SOP model, namely that a CS activates the second inhibitory US process. Rather, it appears to be consistent with a certain feature of real-time models, such as the Sutton and Barto (1981) model or the TD model (Sutton and Barto 1990), namely that a CS can substitute for the US in activating internal reinforcement (see equation 2). However, because both sucrose rewards and reward-predicting odors evoke a long-lasting activation of VUMmx1 (Hammer 1993), a direct implementation of a TD-like reinforcement term via VUMmx1, as suggested by Montague et al. (1995), is contradicted by experimental findings. Because TD computes the first time derivative of the learned prediction of reinforcement \( \gamma Y(t) - Y(t-1) \), presentation of a discrete reinforcement predicting stimulus results in positive reinforcement only for the first time step. How does, therefore, the bee brain compute a predictive error for future rewards and what are putative neural substrates for a US-evoked inhibitory process?

None of the features of reinforcement processing, as postulated by real-time computational models, appear to be directly implemented by the VUMmx1 neuron. The fact that it substitutes the reward in PER conditioning and responds depending on experience nevertheless suggests that a comparison of features of VUMmx1 with reinforcement processing as proposed by theoretical accounts may lead to testable predictions at the behavioral and physiological level. Interestingly, activity of the VUMmx1 neuron is sufficient to mediate the US in olfactory PER conditioning if it is shortly preceded by a CS (forward pairing) but not if a CS is delivered during activation of VUMmx1 (backward pairing) (Hammer 1993). In these experiments, VUMmx1 was activated by current injection for 30 sec, mimicking its long-lasting response to sucrose, and, during backward pairing, the CS was delivered 5 sec after onset of depolarization. Thus, the precise temporal sequence of CS-evoked and VUMmx1 activity controls excitatory conditioning, and activation of VUMmx1 prior to CS presentation must generate a process that prevents conditioning. Therefore, VUMmx1 does not directly implement a positive internal reinforcement signal, because, then, coincidence of CS-evoked and VUMmx1 activity should result in excitatory conditioning. Rather, activation of VUMmx1 should drive a positive reinforcement signal with a fast-decaying time course. Currently, there is no experimental evidence for a neural implementation of the predicted second US-evoked inhibitory process. One possibility is neurons with US-driven long-lasting activity that antagonize the effect of VUMmx1. It is conceivable, however, that activity of VUMmx1 itself results in this second inhibitory process. This would prevent excitatory conditioning if a CS follows the activation of VUMmx1 and could even cause inhibitory learning. In consequence, backward pairing of VUMmx1 activity with an olfactory CS should give rise to inhibitory learning. This notion is also sufficient to account for inhibitory learning with unpaired presentations of CS and US, if one assumes that contextual stimuli, which are associated with reinforcement, activate VUMmx1. Moreover, it makes the strong behavioral prediction that presentation of a neutral stimulus B together with another stimulus A that has been conditioned previously to the US should result in inhibitory learning about B only if B follows A. If B precedes A, this procedure should produce second-order conditioning. Is there, under this assumption, any requirement for the processing of a predictive error within the bee brain? For example, blocking of learning of one component of a binary odor mixture after preconditioning of the other component (Smith and Cobey 1994) has features compatible with the processing of a predictive error for rein-
forcement (Smith 1997). A reduction of the response of VUMmx1 to the sucrose reward if preceded by an odor that has been learned to predict the reward could therefore easily contribute to the blocking phenomenon (for suggestions of how the neural circuit underlying PER conditioning could implement this property, see Hammer and Menzel 1995; Hammer 1997).

Finally, our study supports the notion that the ability of animals to learn both positive and negative predictions is, at least in part, a consequence of the temporal dynamics of the internal processing of events of behavioral significance. In particular, a secondary inhibitory process evoked by reinforcement enables animals to learn in which situations a particular reinforcer will not occur. Because this form of inhibitory learning does not require prior associative experience, both the learning of positive and negative predictions appear to be basic properties of the association process. In contrast, inhibitory learning that depends on prediction errors could have evolved to learn the specific circumstances under which an otherwise expected reinforcer will not occur. The question of whether and how these two forms of inhibitory learning combine to guide experience-dependent behavior requires comparing studies on the rules that govern associative learning with studies on the ecological role of predictive learning. Because olfactory PER conditioning is presumably a form of appetitive learning in foraging behavior in honeybees, it is tempting to speculate on the particular ecological meaning of backward inhibitory learning in this context. Bees foraging in a patch of flowers visit each flower in a short temporal sequence. In some cases, foraging choices seem to be guided by stereotyped motor rules, like "choose the next flower," rather than by the selective use of sensory information (Marden and Waddington 1981; Zimmermann 1982; Kipp 1987). This may lead to erroneous choices. Such errors may have been shortly preceded by visits to rewarding flowers and could therefore result in inhibitory backward conditioning. Inhibitory backward conditioning within a restricted temporal frame would thus facilitate learning of profitable food sources through this experience-dependent replacement of stereotyped motor rules by optimized foraging strategies.

---

Acknowledgments
We thank Martin Giurfa and Tom Carew for helpful comments on the manuscript and Morel Mosolff for his help with experiment 2. This work was supported by a grant from the Deutsche Forschungsgemeinschaft (SPP Physiology and theory of neural nets Me 365/17-2) to M.H.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References


Hellstern et al.


Received September 23, 1997; accepted in revised form December 3, 1997.
Backward inhibitory learning in honeybees: a behavioral analysis of reinforcement processing.

F Hellstern, R Malaka and M Hammer

Learn. Mem. 1998, 4:
Access the most recent version at doi:10.1101/lm.4.5.429

References
This article cites 38 articles, 5 of which can be accessed free at:
http://learnmem.cshlp.org/content/4/5/429.full.html#ref-list-1

License

Email Alerting Service
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.