# From behavioral plasticity to neuronal computation: An investigation of associative learning in the honeybee brain.

Inaugural - Dissertation

to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

Submitted to the Department of Biology, Chemistry, and Pharmacy

Freie Universität Berlin

by

Evren Pamir

from München

January 2013

**The research presented in this dissertation was carried out from August 2009 until January 2013 at the Theoretical Neuroscience & Neuroinformatics group, Freie Universität Berlin, under the supervision of Prof. Dr. Martin P. Nawrot.**

$1^{st}$ Reviewer: Prof. Dr. Martin P. Nawrot - Freie Universität, Berlin
$2^{nd}$ Reviewer: Prof. Dr. Bertram Gerber - Leibniz Institute for Neurobiology, Magdeburg

Date of defense: 08.05.2013

## Acknowledgements

First of all, I would like to thank Martin Nawrot for supervising my PhD project. I am extremely grateful for his continuous support, feedback and ideas. I also thank Martin for creating a friendly and open-minded work atmosphere by putting together the multidisciplinary "neuroinf-all-team". I would like to say THANK YOU to all members of this team, having you as work colleagues over the last years was simply great. In particular, I would like to thank Jan Sölter for knowing the answers to so many questions, Joachim Haenicke for his contributions in joint research ventures, Farzad Farkhooi and Michael Schmuker for their neuroscientific expertise and advice, and Chris Haeusler for constantly proofreading my manuscripts. I also thank Gundula Meckenhäuser, Thomas Rost and Tara Dezhdar for numerous feedback on my project and discussions on neuroscience. I would also like to thank all alumni (Anneke Meyer, Michael Pereira, Jan Meyer, Benjamin Auffarth, Bahadir Kasap) and new members (Louvisa Helgadottir, Stephan Gabler, Rinaldo Betkiewicz, Timo Möller) for the like, and I thank Kathrin Brackwehr for many organizational works.

This theoretical thesis would not have been possible without the work of experimentalists. I would like to express my gratitude to my experimental collaborators from several neurobiology institutes in Germany, who shared their ideas and data with me, and provided invaluable feedback, impulses, and corrections on my work. In particular, I would like to thank Dorothea Eisenhardt, Johannes Felsenburg, Kathrin Marter and Kathrin Gehring (Freie Universität Berlin) for introducing me to the fascinating research on learning and memory in the honeybee, for numerous discussions, and for sharing their data with me. I also thank Randolf Menzel and Neloy Kumar Chakroborty for working together in a side project concerning a cumulative olfactory conditioning paradigm. I am very grateful to Paul Szyszka (University Konstanz) for a truly exciting experimental collaboration on individual learning in the honeybee. I thank Ricarda Scheiner (University Potsdam) for sharing essential experimental data and ideas on learning variability in the honeybee with me. I am grateful to Bertram Gerber and Michael Schleyer (Leibniz Institute for Neurobiology Magdeburg) for working together in a side project on learning in larval Drosophila.

I would like to express a thousand THANKS to Vanessa Casagrande, Margret Franke and Julia Schäffer for taking care of organizing the PhD-Program "Computational Neuroscience" at the Bernstein Center Berlin. I am thankful for an incredible amount of opportunities regarding coursework, exchange of ideas, and financial support provided by this program. I thank the Deutsche Forschungsgemeinschaft and the Federal Ministry of Education and Research for funding this research center as well as for my stipend.

Last but not least, I would like to thank all the other people who made this work possible. I thank Lina, my family, and my friends for moral support, smiles, hugs, and love. I also thank the ever-changing city of Berlin for colorful impressions, daily surprises, and many cheerful moments.

**This dissertation is based on the following three manuscripts:**

**Average group behavior does not represent individual behavior in classical conditioning of the honeybee**

*Authors:* Evren Pamir[1,2], Neloy Kumar Chakroborty[3], Nicola Stollhoff[3], Katrin Barbara Gehring[3], Victoria Antemann[3], Laura Morgenstern[3], Johannes Felsenberg[3], Dorothea Eisenhardt[3], Randolf Menzel[3], and Martin Paul Nawrot[1,2]
*Author contributions:* Research idea, data analysis, modeling, and manuscript by E.P. and M.P.N. Experimental data by N.K.C., L.M., V.A., J.F., K.B.G., N.S., D.E., and R.M.
*Acknowledgments:* We thank Jan Sölter, Bertram Gerber, Björn Brembs, and Paul Szyszka for helpful discussions and Chris Haeusler for comments on the manuscript.
*Manuscript status:* The manuscript has been published. doi: 10.1101/lm.2232711

**Rapid learning dynamics in individual honeybees during classical conditioning**

*Authors:* Evren Pamir[1,2], Paul Szyszka[4], Ricarda Scheiner[5], Martin Paul Nawrot[1,2]
*Author contributions:* E.P. designed the research, analyzed the data, and wrote the paper. P.Z. designed the research, carried out the experiments, analyzed the data and revised the manuscript. R.S. contributed data and revised the manuscript. M.P.N. designed the research and revised the paper.
*Acknowledgments:* We thank Sophie Kroenlein and Nadine Treiber for help with Experiments 1 and 2. We thank Neloy Kumar Chakroborty, Nicola Stollhoff, Katrin Barbara Gehring, Victoria Antemann, Laura Morgenstern, Johannes Felsenberg, Dorothea Eisenhardt, and Randolf Menzel for data contributions. We thank Chris Haeusler and Martin Schwärzel for comments on the manuscript and C. Giovanni Galizia for helpful discussions.
*Manuscript status:* The manuscript has been submitted to the Journal of Neuroscience on January 24, 2013.

**From neuronal computation to behavioral plasticity: A model-based investigation of associative learning in honeybees**

*Authors:* Evren Pamir[1,2], Joachim Haenicke[1,2], Martin Paul Nawrot[1,2]
*Author contributions:* E.P. designed the research, collected and analyzed the data, developed the framework and the model, run the simulations, wrote the paper; J.H. developed the model and revised the paper; M.N. developed the model and revised the paper;
*Acknowledgments:* We thank Michael Schmuker for helpful comments on this work.
*Manuscript status:* A revised version of the current manuscript will be submitted for publication in an international peer reviewed journal.

**Author affiliations:**
1 Bernstein Center for Computational Neuroscience, Berlin.
2 Neuroinformatics & Theoretical Neuroscience, Institute of Biology, Freie Universität Berlin.
3 Institute of Biology - Neurobiology, Freie Universität Berlin.
4 Department of Biology, Universität Konstanz.
5 Institute for Biochemistry and Biology, Zoophysiology, Universität Potsdam.

## Zusammenfassung

Diese Arbeit trägt zu einem besseren Verständnis assoziativen Lernens in der Honigbiene (*Apis mellifera*) bei. Kapitel 2 und 3 dieser Doktorarbeit beinhalten eine statistische Datenanalyse von individuellem Verhalten während klassischer Konditionierung des Proboscisstreckungsreflexes. Es wird gezeigt, dass individuelle Tiere die konditionierte Antwort im Verlauf der Konditionierungsphase schnell und dauerhaft erlernen. Der Gedächtnisabruf in Individuen gleicht einer stetigen Fortsetzung des während der Konditionierung gezeigten Verhaltens. Es wurden keine Anhaltspunkte dafür gefunden, dass ein 24-Stundengedächtnis, welches durch mehrere Konditionierungstrials induziert wurde, stärker oder spezifischer ist als ein durch einen Trial induziertes Gedächtnis. Diese Ergebnisse erweitern das gegenwärtige Lern- und Gedächtnismodell in der Honigbiene, welches auf Basis der Analyse von Gruppenmittelwerten entwickelt wurde. Die Analyse von Verhaltensdaten liefert zudem die Grundlage für ein erneuertes theoretisches Bild von assoziativem Lernen in der Honigbiene. In Kapitel 2 und 3 wird die Rescorla-Wagner Theorie erweitert und auf dem Niveau von Einzeltieren angewendet. Kapitel 4 stellt ein computationales Modell vor, anhand dessen sich die Auswirkungen von externen Stimuluseigenschaften, interner Enkodierung und Prinzipien neuronaler Computation auf den zeitlichen Verlauf von Verhaltensänderungen während klassischer Konditionierung in einer Vielzahl unterschiedlicher Lernaufgaben untersuchen lassen. Kapitel 4 liefert damit die Grundlage für die Zusammenführung von Verhaltensdaten und neurophysiologischen Daten zu assoziativem Lernen in der Honigbiene.

## Summary

This research contributes to a better understanding of learning and memory in the honeybee (*Apis mellifera*). Chapters 2 and 3 of this thesis include a statistical data analysis of individual behavior during classical conditioning of the proboscis extension response. It is shown that individual honeybees rapidly acquire a stable conditioned response during the conditioning phase. Memory retention in individuals resembles a steady continuation of the behavior expressed during training. No behavioral evidence is found for a more stable or specific 24-hour-memory after multiple-trial than after single-trial conditioning. These results extend the current model on learning and memory in the honeybee, which has been established at the group-average level of analysis. The behavioral data analysis also provides the basis for a renewed theoretical account of associative learning in the honeybee. In chapters 2 and 3, the Rescorla-Wagner theory of associative learning is extended and applied at the level of individuals. Chapter 4 introduces a computational model for investigating the effect of external stimulus properties, internal encoding schemes, and principles of neuronal computation on the dynamics of behavioral plasticity during classical conditioning in a variety of different learning tasks. By this, chapter 4 introduces a computational framework for the integration of behavioral and neurophysiological data on associative learning in the honeybee.

**Keywords:**
behavioral neuroscience, classical conditioning, insect brain, honeybee, associative learning, plasticity, neuronal computation, computational modeling

# Contents

# Chapter 1

# General introduction

## 1.1 Statement of the problem

The behavioral study of associative learning under controlled laboratory conditions in vertebrate and invertebrate animal models has greatly advanced our understanding of the nervous system. The long-term goal of this research lies in explaining behavior by neuronal functioning. However, behavior only provides an indirect measure of the state of the brain. The question arises, *what* and *how* can we learn about the brain when monitoring behavior?

To date, one problem lies in the poor correspondence between the found properties of associative learning observed in behavior, and neuronal plasticity measured in the brain, which hinders the attempt to explain the first by the latter (Gallistel and Matzel, 2012). For example, the induction of long-term-potentiation, a well-studied form of synaptic plasticity, requires several repetitions of a stimulation protocol (Feldman, 2012). On the other hand, learning at the behavioral level is often complete within a single training trial (see Table 1 in Gallistel and Matzel, 2012, for a list of disparate properties of long-term-potentiation and associative learning). To point to another problem, experiments on animal learning typically make use of group-average performance scores in order to evaluate the temporal evolution of learning across training trials. However, Gallistel et al. (2004) showed that group-average performance did not capture the characteristics of individual learning for several vertebrate learning paradigms. As a consequence, the search for the neuronal correlates of learning in the individual brain may be misinformed. To address a third problem, Carandini (2012) pointed out that explaining behavior by neuronal processes may be impossible without first establishing an understanding of the computations performed by the neurons or neuronal circuits under investigation. Carandini argues that "researchers of circuits and of behavior go furthest when they speak a common language of computation".

This thesis tries to refine and extend the current understanding of behavioral plasticity during classical conditioning in the honeybee (*Apis mellifera*), which in turn may contribute to a better understanding of neuronal plasticity in this animal model. Chapters 2 and 3 of this

thesis present a detailed analysis of behavior in order to establish a reference for the dynamics of associative learning in individual honeybees under various different training conditions. Chapter 3 explains how the analysis of individual behavior can provide more informative constraints for molecular hypotheses on memory consolidation. Finally, Chapter 4 follows the suggestion by Carandini (2012) and implements a computational model of associative learning which may build a bridge between findings on behavioral plasticity on one hand, and neuronal processes in the honeybee brain on the other.

## 1.2 A canonical assay for associative learning: Classical conditioning of the proboscis extension response in the honeybee

Classical conditioning assays provide an invaluable tool for the investigation of learning and memory, both at the behavioral and physiological level. Current research in the honeybee now builds on more than fifty years of experience in training honeybees to learn contingencies between sensory stimuli and gustatory reinforcement: classical conditioning of the proboscis extension response (PER) was first introduced by Kimihisa Takeda in 1961 (Takeda, 1961), and since then has undergone continuous improvement and alterations in scope (Giurfa and Sandoz, 2012). Today, PER conditioning in honeybees is a canonical paradigm enabling the parallel investigation of learning *and* of its neuronal correlates (Sandoz, 2011). Numerous comprehensive reviews describe the details of the experimental procedure (Frost et al., 2012; Matsumoto et al., 2012; Giovanni Galizia, 2011), and video material showing the training of harnessed honeybees in the laboratory is available online (Felsenberg et al., 2011; www.jove.com/video/2282).

The most basic version of this Pavlovian conditioning paradigm makes use of only one olfactory stimulus during training, which is referred to as absolute conditioning. In the conditioning phase, individual harnessed honeybees receive several paired presentations of an odorant (conditioned stimulus CS) and a sucrose reward (unconditioned stimulus US). Typically, the CS onset precedes the US onset by a few seconds, such that both stimuli overlap for 1 or 2 seconds. The US is delivered to the honeybee by the experimenter by first touching the antennae with a toothpick (or needle) soaked in sucrose solution and than delivering a few droplets of sugar solution to the proboscis (see for example Matsumoto et al., 2012). The conditioned response (CR) of an individual, namely the proboscis extension or non-extension prior to the onset of the US, is visually recorded by the experimenter in a binary format. Measuring the average probability of a proboscis extension in a group of honeybees across several conditioning trials results in the acquisition function, or learning curve. This curve provides a measure of learning or behavioral plasticity during training at the group-average level. Analogous, the group-average CR probability after training, for example at 24 hours, provides a measure of memory retention (Menzel, 1990). The vast majority of behavioral

studies on learning and memory in the honeybee has so far followed this methodology, which is based on averaging binary conditioned responses. Some studies showed that more subtle differences in learning can be found in non-binary measures of the proboscis extension. Smith and Menzel (1989a,b) determined the duration and onset latency of the PER by high-speed video recordings, as well as by measuring the electromyogram of the M17 muscle, which is one of the muscles of the mouthparts involved in the proboscis extension. Absolute conditioning experiments sometimes also compare response levels between the conditioned and a novel odor in the retention test (Biergans et al., 2012; Matsumoto et al., 2012), which allows the separation of non-associative components of the conditioning procedure from the purely associative ones (Tully, 1984).

Although only one odorant is involved during the training phase, absolute conditioning already spans a large experimental parameter space for possible investigations. The effect of several parameters on learning and memory have already been described in the experimental literature, such as the number of conditioning trials (Menzel, 1990; Menzel et al., 2001), the inter-trial-interval between conditioning trials (Menzel et al., 2001), the inter-stimulus-interval between CS and US (Szyszka et al., 2011), the concentration of the conditioned odorant or the concentration of the sucrose solution employed as reinforcement (Bitterman et al., 1983; Scheiner et al., 2005). Early work in the honeybee showed that honeybees rapidly learn during absolute conditioning: typically $30\% - 80\%$ of all animals acquire the CR after the first conditioning trial (Bitterman et al., 1983). To make the task more difficult for the animal, researchers moved to non-absolute (or non-elemental) conditioning protocols from experimental psychology (Rudy and Sutherland, 1992, 1995). Surprisingly, honeybees can cope extremely well with these more demanding or "cognitive" learning tasks (Giurfa, 2003; Menzel et al., 2007). For example, in a negative patterning protocol honeybees have to learn that two olfactory stimuli presented together are not followed by reinforcement, whereas individual stimuli alone are rewarded. Unlike other insect species such as the fruit fly (Young et al., 2011), honeybees are able to learn this task (Deisig et al., 2001, 2002). Several other non-absolute protocols have been tested in honeybees, such as positive patterning (Deisig et al., 2001), learning of ternary and quaternary mixtures (Deisig et al., 2003), blocking (Gerber et al., 1998; Guerrieri et al., 2005), and biconditional discrimination (Chandra and Smith, 1998; Hellstern et al., 1998).

As outlined above, the study of learning and memory in the honeybee has mainly followed the methodology of group-average learning curves and retention scores. Consequently, behavioral plasticity in absolute and non-absolute PER conditioning has been predominantly described at the group-average level. However, as demonstrated by Gallistel et al. (2004) (see above), behavioral plasticity at the group-average level does not necessarily coincide with the learning dynamics in individuals. Several insightful features of individual learning may be missed by only relying on the group-average perspective. Interestingly, for the honeybee,

several sources of learning variability in individuals have already been described, such as behavioral role (Scheiner et al., 1999; Scheiner and Malun, 2001; Scheiner et al., 2003), age (Behrends and Scheiner, 2012), or genetic factors (Dukas, 2008; Liang et al., 2012). Scheiner et al. (2005) showed that measuring the sensitivity to sucrose in individuals prior to conditioning provides a possibility to uncover and calibrate inter-individual differences in learning performance. The authors also noticed that this pre-assessment of individuals may have important consequences for the correct comparison of group-average memory retention in different experimental groups. However, to date only little emphasis has been put on a refinement of methodology in this respect (Matsumoto et al., 2012).

Chapters 2 and 3 of this thesis provide a detailed analysis of inter-individual learning differences during PER conditioning in absolute and non-absolute conditioning tasks. This analysis shows how current hypotheses about the effect of training parameters on memory retention can be refined. The analysis of Chapters 2 and 3 also provides the basis for the modeling work presented in Chapter 4, which tries to explain the dynamics of behavioral plasticity by principles of neuronal computation in the honeybee brain circuitry.

## 1.3 Computational modeling of associative learning in the insect brain

The massive amount of data on behavioral plasticity in the honeybee obtained in absolute and non-absolute conditioning experiments naturally calls for a comprehensive integration by computational theories or models. Models of associative learning come in different levels of biological detail. The ultimate goal of any model approach consists in a better understanding of a specific aspect of a complex problem (Gluck and Myers, 2001), and hence more detailed models cannot be per se considered superior to simpler ones (Dayan and Abbott, 2001). The Rescorla-Wagner model (Rescorla and Wagner, 1972) is probably one of the most famous "simple" models from associative learning theory. In its essence, this model assumes that learning is driven by prediction errors, a hypothesis which has found widespread use in the computational neuroscience community (Gluck and Myers, 2001). In the second and third chapter of this thesis I will explain how this model can be fit to behavioral data in the honeybee, while taking into account inter-individual differences in learning performance.

Descriptive models of the Rescorla-Wagner type make quantitative predictions for the temporal evolution of learning as a function of stimulus-reward contingencies as defined by a given conditioning protocol. This type of model has been successfully applied to conceptualize behavioral findings on non-elemental learning in the honeybee (Deisig et al., 2003) and the fruit fly (Young et al., 2011). While these models provide insights into the computations performed by the brain at an abstract level, they do not touch the issue of how these computations may be implemented in the biological system.

Olfactory stimulus processing and associative learning in the honeybee has been characterized by numerous physiological studies (Sandoz, 2011; see Table 2 in Himmelreich and Grünewald, 2012, for a list of learning-related events in the honeybee brain). The entire honeybee brain spans less than a cubic millimeter in volume and comprises about one million neurons. Computational studies in insects typically concentrate on a subset of neurons from the whole sensor-to-motor circuitry, which are believed to implement functions such as stimulus detection and preprocessing (Schmuker et al., 2011; Assisi et al., 2012; Locatelli et al., 2012), or high-order stimulus representation and associative learning (Smith et al., 2008; Huerta and Nowotny, 2009; Wessnitzer et al., 2012; Smith et al., 2012; Arena et al., 2012; Huerta, 2013). In the honeybee brain, olfactory stimuli are detected by olfactory receptor neurons in the antennae (see Sandoz, 2011, for a comprehensive review on the honeybee olfactory pathway). Olfactory receptor neurons project to 165 olfactory glomeruli in the antennal lobe, the primary olfactory center in the honeybee brain. Each glomerulus is assumed to receive information from only one receptor type in the ORNs. Odor identities and concentrations are represented by spatio-temporal activation patterns in these glomeruli (Galizia et al., 1999). The Antennal lobe circuit, which consists of local interneurons and projection neurons, implements several stimulus preprocessing steps, such as gain control and lateral inhibition, which provide a better discrimination of odor concentrations or identities, respectively (Schmuker et al., 2011; Assisi et al., 2012). Refined olfactory information from this circuitry is relayed to higher brain centers such as the mushroom body and the lateral horn by different types of projection neurons (Sandoz, 2011). The intrinsic neurons of the mushroom body, the Kenyon cells, then converge on a small number of extrinsic neurons, which in turn are believed to target pre-motor centers that drive the extension of the proboscis.

The simplified sensor-to-motor circuit described here has been the focus of several computational studies (Huerta and Nowotny, 2009; Wessnitzer et al., 2012; Smith et al., 2012). From a theoretical viewpoint, the architecture of the insect brain makes the mushroom body an ideal location for a high-dimensional sparse representation of combinatorial activation patterns from the antennal lobe. Moreover, this enables behaviorally relevant values of these high-order representations to be easily learned by some type of reward-dependent synaptic plasticity at the mushroom body output. This idea complies well with the current working hypotheses on associative learning in biology, which locates the olfactory memory trace at the mushroom body output (Heisenberg, 2003; Gerber et al., 2004). Importantly, appetitive reinforcement during associative learning has been shown to be provided by a giant octopaminergic neuron (Hammer, 1993), similar to dopaminergic neurons in the mammalian brain (Schultz, 2006). Based on these ingredients, namely a sparse odor representation by Kenyon cell firing and a teacher signal at the mushroom body output, several different theoretical accounts of synaptic plasticity have been proposed in the insect brain. Non-spiking models make use of so-called three-factor learning rules, in which weight changes depend on pre- and post-synaptic activity

(Hebb, 1949), as well as on a teacher signal (Huerta and Nowotny, 2009; Smith et al., 2012). Spiking implementations, on the other hand, make use of spike-timing-dependent plasticity (Smith et al., 2008; Wessnitzer et al., 2012), as observed in the mammalian brain (Caporale and Dan, 2008; Feldman, 2012).

Computational work in the honeybee cannot draw from direct experimental evidence of spike-timing-dependent plasticity at the mushroom body output (in contrast to for example work in the locust by Cassenaer and Laurent, 2012). In Chapter 4 of this thesis, I will present a computational model of associative learning in the honeybee at a level of detail comparable to the study by Huerta and Nowotny (2009), hence without an implementation of a neuronal spiking mechanism. The model intends to explain the temporal evolution of learning during absolute and non-absolute conditioning by a simple stimulus processing and associative learning scheme along the sensor-to-motor pathway in the honeybee. In my model approach I investigate how stimulus properties, mushroom body encoding, and neuronal computations such as different associative learning rules or divisive normalization affect behavioral plasticity. As suggested by Carandini (2012), this chosen level of model complexity may provide a bridge between biological details on neuronal functioning on the one side and the emergent properties of behavior on the other.

# Chapter 2

# Average group behavior does not represent individual behavior in classical conditioning of the honeybee

**Abstract**

Conditioned behavior as observed during classical conditioning in a group of identically treated animals provides insights into the physiological process of learning and memory formation. However, several studies in vertebrates found a remarkable difference between the group-average behavioral performance and the behavioral characteristics of individual animals. Here, we analyzed a large number of data (1640 animals) on olfactory conditioning in the honeybee (*Apis mellifera*). The data acquired during absolute and differential classical conditioning differed with respect to the number of conditioning trials, the conditioned odors, the inter-trial intervals, and the time of retention tests. We further investigated data in which animals were tested for spontaneous recovery from extinction. In all data sets we found that the gradually increasing group-average learning curve did not adequately represent the behavior of individual animals. Individual behavior was characterized by a rapid and stable acquisition of the conditioned response (CR), as well as by a rapid and stable cessation of the CR following unrewarded stimuli. In addition, we present and evaluate different model hypotheses on how honeybees form associations during classical conditioning by implementing a gradual learning process on the one hand and an all-or-none learning process on the other hand. In summary, our findings advise that individual behavior should be recognized as a meaningful predictor for the internal state of a honeybee - irrespective of the group-average behavioral performance.

## 2.1 Introduction

Learning and memory formation in vertebrates and invertebrates have been studied on the basis of a large range of classical and operant conditioning paradigms. Typically, the interpretation of experimental results relies on performance measures that were derived by averaging over behavioral observations from identically treated animals. However, several studies have recognized the inadequacy of group-average measures to capture the characteristics of individual behavior and, consequently, the learning- induced changes in individual brains (Krechevsky, 1932; Restle, 1965; Hanson and Killeen, 1981; Estes, 2002; Brown and Heathcote, 2003; Cousineau et al., 2003). Most notably, Gallistel et al. (2004) found that the gradually increasing learning curve observed in many vertebrate learning paradigms reflected an artifact of group averaging. The behavioral performance of individuals appeared to be characterized by an abrupt and often step-like increase in the level of response.

To our knowledge and in contrast to the vertebrate literature (see Gallistel et al., 2004), surprisingly little is known of a possibly heterogeneous expression of behavior for the most frequently applied invertebrate conditioning paradigms. For the fruit fly (Drosophila melanogaster) it appears to be common sense that the group-average behavioral measures adequately represent the probabilistic expression of behavior in individuals - a notion that goes back to an early study by Quinn et al. (1974). In the following, we focus on a classical conditioning paradigm in the honeybee (*Apis mellifera*) - the reward-based olfactory conditioning of the proboscis extension response (Takeda, 1961; Bitterman et al., 1983). In this paradigm, a group of harnessed honeybees is individually trained by forward pairings of an olfactory stimulus (CS) with a sucrose reward (US). The conditioned response (CR), namely, the extension or nonextension of the proboscis at each of these conditioning trials, is typically documented in a binary form. Additional characteristics of the proboscis extension can be captured by using a high-speed video or by recording the electromyogram of the muscles involved in the proboscis extension (Rehder, 1987; Smith and Menzel, 1989a,b). Importantly, conditioning of the proboscis extension response allows one to simultaneously measure brain activity by means of electrophysiology or calcium imaging (Giurfa, 2007; Menzel et al., 2007), giving access to the neuronal correlates of learning and memory. The molecular mechanisms of memory consolidation in the honeybee brain can be studied by combining classical conditioning with in vivo pharmacological interventions or postmortem biochemical analysis (Stollhoff et al., 2005; Eisenhardt, 2006; Schwärzel and Müller, 2006).

In the present study, we determine the behavioral characteristics of individual honeybees during classical conditioning on the basis of a large collection of experimental data. Specifically, we ask: (1) How well does the group-average behavioral performance represent the behavioral performance of individuals during absolute conditioning, differential conditioning, and extinction? (2) How do individual animals learn during absolute conditioning, and how does this learning translate into behavior? The Results section is divided into two parts. In

the first part, we answer the first question by means of an exploratory data analysis. In the second part, we consider three different answers to the second question by evaluating the eligibility of three different generative models. The implications of our findings for the analysis and interpretation of behavioral and physiological data from the honeybee are discussed.

## 2.2 Results

### 2.2.1 Experimental data

We analyzed data from a number of independent experiments that were conducted at the Institute of Biology - Neurobiology of the Freie Universität Berlin between the years 1999 and 2009. None of the experiments were originally designed for the purpose of the present study. We collected a total of 17 data sets comprising 1640 animals (see Table 2.1). The animals in data sets 1-15 were trained by using an absolute classical conditioning protocol, while the animals from the data sets 16 and 17 were differentially conditioned (see Materials and Methods). The data sets differ with respect to the number of conditioning trials $m$, the temporal inter- trial interval during acquisition (ITI), the time-point of the memory retention test ($T$), and the odor that was used as the conditioned stimulus (CS). A subset of animals ($n = 217$) from data set 1 was subjected to an extinction protocol.

### 2.2.2 For absolute conditioning, the group CR probability does not represent individual behavior

For each of the data sets 1-15 (absolute classical conditioning), we asked how well the group CR probability $P(x_t = 1)$ in a given trial $t$ represented the CR probability of subgroups of animals. We compared the behavioral performances of two disjoint subgroups defined by their response in the previous trial ($x_{t-1} = 1$ or $x_{t-1} = 0$). Figure 2.1 A shows an example of the standard data analysis (see Materials and Methods) for the data set 10. While performing this analysis for data sets 1-15 we found a remarkably uniform pattern during conditioning and in the memory retention test (Fig. 2.1 B,C): Animals that responded in the previous trial $t-1$ always had a higher chance for responding in trial $t$ than animals that did not respond in trial $t-1$. The differences between the subgroup CR probabilities $P(x_t = 1|x_{t-1} = 1)$ and $P(x_t = 1|x_{t-1} = 0)$ were statistically significant ($\chi^2$ test, $\alpha = 0.05$) in 39 out of 39 cases during conditioning (Fig. 2.2 B), and in 15 out of 20 cases during memory retention (Fig. 2.2 C). For each of the 15 data sets we also computed the mean probabilities for the observations $(x_t = 1|x_{t-1} = 0)$ and $(x_t = 1|x_{t-1} = 1)$ across all possible trials. Computing the average of these mean probabilities across data sets 1-15 yielded ($\pm$ SD): $P(x_t = 1|x_{t-1} = 1) = 0.98 \pm 0.05$ and $P(x_t = 1|x_{t-1} = 0) = 0.43 \pm 0.14$.

### 2.2.3 Differential conditioning is consistent with absolute conditioning

We next asked whether the results obtained for absolute conditioning also applied to data from differential conditioning experiments. For the data sets 16 and 17 we separately analyzed the CS+ and CS- trials using the standard data analysis (see Materials and Methods). We found that animals that responded to the CS+ (CS-) in the previous trial $t-1$ ($t'-1$) had a higher chance for responding to the CS+ (CS-) in trial $t$ ($t'$) than animals that did not

respond to the CS+ (CS-) in the previous trial (Fig. 2.2). Once animals responded to the CS+, or once animals did not respond to the CS- anymore, the response probability remained rather constant in the following trials. Thus, the group CR probability does not adequately represent the behavioral characteristics of individual animals during differential conditioning.

### 2.2.4 The cessation of the CR is abrupt in individuals during extinction

A subset of animals ($n = 217$) from data set 1 was presented in five extinction trials 24 h after conditioning (Fig. 2.3 A,B). After the extinction session, the animals were divided into five groups, each of which was tested for spontaneous recovery from extinction at a different time point (Fig. 2.3 C; see also Fig. 2.3 in the original study by Stollhoff et al., 2005). By using the standard data analysis (see Materials and Methods) we asked whether the heterogeneous expression of behavior observed during conditioning was also present during extinction and subsequent memory retention. For all trials, we again found that animals that responded in the previous trial $t-1$ had a higher chance for responding in the trial $t$ than animals that did not respond in the previous trial (Fig. 2.3 A-C). The analysis also revealed that once individual animals had ceased to respond during extinction ($t = 4-8$), they had a high chance of remaining nonresponding in subsequent extinction trials (Fig. 2.3 B). Hence, analogous to our findings for the acquisition of a CR during conditioning, the cessation of the CR during extinction is characterized by an abrupt rather than gradual change of CR probability in individuals.

On the basis of our previous analysis, we hypothesized that the presence or absence of the CR in the retention test ($t = m + 1$) may serve as a binary indicator for the actual learning success of individual animals. Consequently, showing a CR on this trial ($t = 4$) may indicate that an associative memory has been induced by the conditioning procedure, while not showing a CR on this trial may indicate that no association or only a poor association has been formed. To test our hypothesis we divided the 217 animals into two disjoint subgroups that were defined by their response on trial $t = 4$ ($x_4 = 1$ and $x_4 = 0$). As expected, our heuristic selection criterion yielded a steady performance difference between the two subgroups during conditioning (Fig. 2.3 D): the animals in the first subgroup ($x_4 = 1$) had a higher response probability than the animals in the second subgroup ($x_4 = 0$). This performance difference also persisted during extinction and memory retention (Fig. 2.3 E,F), which may provide evidence for the heterogeneity of the 217 conditioned animals with respect to the associative memories formed. We tested the differences of the two subgroups for statistical significance by a $\chi^2$ test and found that the differences were always significant during acquisition ($t = 2, 3$) and extinction ($t = 5 - 8$), but only significant once in the subgroups tested for memory retention (25 h). It should be noted that the test power was much lower in the retention tests because the group sizes were reduced by a factor of 4.

### 2.2.5 The simple learning-curve model must be rejected

Given the typical asymptotic rise of the group CR probability in a classical conditioning experiment, one may assume that the shape of the curve reflects a gradual increase of associative strength $AS(t)$ across trials in individuals. One may further assume that all animals are identical with respect to this learning process and that individual animals express a CR at trial $t$ with probability $P^{LCM1}(x_t = 1) = AS(t)$. We tested the eligibility of this model hypothesis, termed the simple learning-curve model, on data sets 1-15 (see Fig. 2.4 A and Materials and Methods for the full model description). While this model fit the group CR probabilities very accurately (Fig. 2.6 A, below), it did not capture the CR probabilities at the level of subgroups. As we expected from our data analysis in the previous sections, the empirical subgroup probabilities $P(x_t = 1|x_{t-1} = 1)$ and $P(x_t = 1|x_{t-1} = 0)$ significantly deviated (two-tailed binomial test, $\alpha = 0.05$) from the model predictions in 89 out of 104 cases (Fig. 2.5 A).

### 2.2.6 The learning-curve model can be saved by two additional features

In this section we maintain the previous notion of a gradual increase of associative strength in individuals across conditioning trials; however, we add the possibility that the rate of learning may not be the same in all individuals. Furthermore, we use a simple performance rule, effectively a threshold criterion, which maps the associative strength in individuals to the CR probability (see Fig. 2.4 B and Materials and Methods for a full description of this model). By fitting this extended learning-curve model to data sets 1-15 we found that it provided a possible explanation for the observed behavioral sequences. The empirical subgroup probabilities $P(x_t = 1|x_{t-1} = 1)$ and $P(x_t = 1|x_{t-1} = 0)$ deviated significantly from the model predictions in only nine out of 104 cases (Fig. 2.5 B). We also noticed that the estimated probability distributions were skewed to the highest learning-rate interval (see Fig. 2.7, below). Computing an average probability distribution over data sets 1-9, we found that the highest resolvable learning-rate interval is adopted by 48% of the animals, while the second, third, and fourth highest intervals are adopted by 26%, 18%, and 8% of the animals, respectively. For data sets 10-15, in which the lower learning-rate intervals were resolvable, we always found another peak of the distribution at the respective lowest interval.

### 2.2.7 A two-state hidden Markov model captures the behavioral sequences

Our data analysis revealed that once animals were responding in a given trial, they had a very high probability of again responding in the next trial. We hypothesized that this type of response behavior could be adequately captured by a hidden Markov model with two hidden states, $k = 1$ and $k = 2$, and two possible observations, $x_t = 0$ and $x_t = 1$ (see Fig. 2.4 C

and Materials and Methods for a full description of this model). Fitting the two-state hidden Markov model to data sets 1-15, we found that the parameter estimation yielded highly similar results for each of the data sets. Furthermore, for each data set the estimated parameters did not depend on the chosen starting values. We report here their mean values averaged over data sets 1-15 (±SD): $P(k_1 = 1) = 1.00$, $P(k_t = 2|k_{t-1} = 1) = 0.43 \pm 0.16$, $P(k_t = 2|k_{t-1} = 2) = 0.96 \pm 0.04$, $P(x_t = 0|k_t = 1) = 1.00 \pm 0.01$, $P(x_t = 1|k_t = 2) = 0.94 \pm 0.04$. These values express the following prototypic scenario: At $t = 1$ all animals are in the state $k = 1$. Animals in this state have a very low probability for showing a CR and a moderate probability for making a transition into the state $k = 2$. The second state is characterized by a very high probability for showing a CR and a very high probability for remaining in this state. As a consequence, once an individual animal has made a transition to the second state $k = 2$, it will show a stable expression of the CR in the subsequent trials. As in the case of the extended learning-curve model, we found that the hidden Markov model provided a possible explanation for the observed behavioral sequences. The empirical subgroup probabilities $P(x_t = 1|x_{t-1} = 1)$ and $P(x_t = 1|x_{t-1} = 0)$ only significantly deviated from the model predictions in 11 out of 104 cases (Fig. 2.5 C).

### 2.2.8 Model comparison by cross-validation

The more complex an explanation, the better it can capture a given set of data. Simple explanations, however, are often the best. In terms of complexity, the simple learning-curve model is the simplest, comprising only two free parameters. In turn, the extended learning-curve model is the most complex because it uses an unconstrained empirical probability estimate for the learning-rate distribution. For a data set with $L$ trials, the extended learning model has $L + 1$ free parameters. With four free parameters, the hidden Markov model lies between the other two models in terms of complexity. By fitting the three models to data sets 1-15 we found that the extended learning-curve model and the hidden Markov model could capture the behavioral observations, while the simple learning-curve model could not (Figs. 2.5, 2.6). However, this finding may result from their higher complexity rather than from their higher explanatory power. To directly compare the eligibility of the three models, we computed their log likelihoods by a cross-validation algorithm for each of the data sets 1-15 (see Materials and Methods). This procedure confirmed that the simple learning-curve model had to be rejected, since it performed worse than each of the other two models in all cases (see Table 2.2). We also found that the extended learning-curve model was more likely than the hidden Markov model in 12 out of 15 cases.

## 2.3 Discussion

**The group CR probability is a poor estimate for the behavior of individual honeybees** Gallistel et al. (2004) analyzed several vertebrate learning paradigms for how well the group-average performance measures represented the behavioral performance of individuals. They found that the gradual and negatively accelerated increase seen at the population level defining the empirical learning curve was an artifact of group averaging. Individual behavior was characterized by an abrupt and often step-like change in the level of responding. Since at each time point the population comprises two types of animals, those that have acquired the CR, and those that have not, the investigators suggest describing the experimental data by individual onset latencies rather than by group-average measures.

As we have shown, the case is very similar for classical conditioning of the proboscis extension response in the honeybee. For both absolute and differential conditioning, we found a very high probability that a bee would again extend its proboscis, given that it had done so in the previous trial. This high probability clearly excludes the adequacy of using the gradually increasing group CR probability to represent the response probability of individuals during the training phase. Most importantly, we found that the heterogeneous expression of behavior observed during the conditioning phase persisted during memory retention, which indicates a heterogeneity in a group of identically treated animals with respect to long-term memory formation. Analyzing data in which animals were tested for spontaneous recovery from extinction provided additional evidence for this notion. Interestingly, we also observed a rapid and stable change of the response probabilities when honeybees had to learn that stimuli were not followed by a reward, as was the case for the unrewarded stimuli presentations during differential conditioning and extinction. Thus, once individual animals did not respond to an unrewarded stimulus in a given trial, they had a high probability of not responding in subsequent trials. In summary, individual behavior is characterized by abrupt and stable changes in response probabilities, contrary to the gradual changes in group CR probability observed at the population level.

**How representative is the mean in the fruit fly?** Our results differ from the commonly held notion of a homogeneous expression of conditioned behavior in the fruit fly, as first reported in a study by Quinn et al. (1974). For a binary choice olfactory-conditioning paradigm, Quinn et al. (1974) asked whether the group-average performance observed after conditioning arose from some heterogeneity in the population, or whether it was due to a stochastic component in the behavior of all of the flies. To answer this question, the investigators separated the flies that made the correct choice from those that did not, and 24 h later retrained and retested each group. Since the performance of both groups was the same, the investigators concluded that the expression of behavior was probabilistic in each fly, and that there was no evidence for an intelligent subset of the population. A recent study by Chabaud et al. (2010) made

a more ambiguous observation when tracking the choice behavior of individual fruit flies during memory retention. For two types of training protocols, the investigators found that the expression of behavior was probabilistic during the test, because the choice of individual flies at the end of the test was not determined by their first choice in the test. However, for a third training protocol, the final memory score was determined by the first choice, yielding a bipolar distribution of the individual memory scores. It remains to be more thoroughly investigated whether serial correlations in the behavior of individuals during training, and between training and testing as reported here, is a specific result for the honeybee or whether it also applies to other invertebrates.

**How do individual honeybees learn?** For the honeybee, several experimental parameters and conditions have been described as having a decisive effect on the strength, specificity, and stability of associative memories by com- paring group-average CR probabilities in differently treated groups (for the number of conditioning trials and intertrial inter- vals, see Bitterman et al., 1983; Sandoz et al., 1995; Gerber et al., 1998; Menzel, 1999, 2001; for bee age or season see Behrends and Scheiner, 2010; Hadar and Menzel, 2010; for US strength, see Menzel, 2001; Scheiner et al., 2004). To pick the most basic experimental parameter, if one group A is trained with more conditioning trials than another group B, this typically yields a higher group CR probability in group A at the end of the conditioning phase, as well as in the memory-retention test compared with group B. The classical interpretation of this canonical observation is that more training yields stronger memories, implicitly assuming that each of the two groups is rather homogenous with respect to learning and memory formation under given experimental conditions. However, both of the extended learning-curve model and the hidden Markov model are at odds with this line of reasoning.

According to the extended learning-curve model, the population is heterogeneous with respect to learning rates (see Fig. 2.7). Consequently, in any trial $t$, the population will comprise animals with associative strengths below threshold that do not show a CR, and animals with associative strengths above threshold that show a CR with high probability. Hence, the group CR probability reflects the ratio between the sizes of the two subsets of animals at any trial $t$. If group A is trained with more trials than group B, then this results in a higher group CR probability at the end of the conditioning phase as well as in the memory retention test in group A, however, not because stronger memories have been induced in individuals of group A, but simply because the ratio between the sizes of the two subsets has been shifted to a larger value. To be more specific, the extended learning-curve model assumes that animals with an associative strength above threshold increase their associative strength with further conditioning trials, but this process does not contribute to the shift of the observed ratio, nor is it visible in the individual behavioral sequences. To conclude that associations in individuals have been strengthened, would at least require nonbinary behavioral measures. For example, when recording the potentials of the M17 muscle involved

in the proboscis extension response, Smith and Menzel (1989a,b) observed a gradual rise of the muscle potentials and a gradual decrease of the response latencies with the number of conditioning trials.

The hidden Markov model assumes all-or-none learning, which is that at each trial $t$, a given animal has a certain success rate $P(k = 2|k = 1)$ for learning a stimulus-reward association in an all-or-none fashion. Whenever an association has been learned by an animal, this association is highly stable across the remaining trials. According to this view, more training trials increase the number of success opportunities for each animal and, consequently, more training trials will increase the number of animals in the second state. At each trial $t$, the group CR probability then simply reflects the ratio between the number of animals in the naive state ($k = 1$) and in the learned state ($k = 2$); however, the group CR probability does not contain any information about the state of an individual animal. (There are two exceptions: when all animals occupy either the first or the second state). It should be noted that in the hidden Markov model view, individual animals only need one successful trial to establish a stable association. Some experimental support for this possibility comes from a study that showed honeybees can form a stable and longlasting memory even after a single conditioning trial (Sandoz et al., 1995). One can further hypothesize that the success rate $P(k = 2|k = 1)$ depends on a large number of experimental conditions, as well as on intrinsic conditions of the respective animal such as hunger, motivation, hormonal status, health, or age, while the probability for maintaining a formed association is possibly invariant to these factors.

To summarize, the extended learning-curve model and the hidden Markov model represent two alternative hypotheses for the dynamics of associative learning in the honeybee during classical conditioning. Further experiments to test the various consequences of the two hypotheses are underway.

**Implications for data analysis** Our findings advise that individual behavior should be recognized as a meaningful predictor for the internal state of a honeybee - irrespective of the group CR probability. In particular, this suggests that the analysis of parallel behavioral and physiological recordings should be carried out at the level of individual animals. Several studies have demonstrated how this can lead to a more informative analysis. For a rule-learning task in rats, Durstewitz et al. (2010) found that neuronal activity recorded from the prefrontal cortex was in tight temporal relation to behavioral performance shifts in individuals. For the honeybee, two recent studies divided groups of identically treated animals into subgroups of so-called learners and nonlearners on the basis of a heuristic behavioral selection criterion (Roussel et al., 2010; Rath et al., 2011). Both studies then found differences in the simultaneously recorded CA-imaging signals between the two subgroups. Here, we also suggest that behavioral studies that use in vivo pharmacological interventions or post-mortem biochemical analysis of honeybee brains should take into account individual behavior as a possibly discriminative

factor.

Our analysis offers several possibilities to illustrate the heterogeneity in a given set of behavioral data. Adopting the extended learning-curve model, the distribution of learning rates across individuals can be empirically estimated and visualized. Adopting the hidden Markov model, the transition and observation probabilities for the naive and the conditioned state can be determined. Under a model-free perspective, conditional probabilities can be computed as demonstrated in the exploratory data analysis part of this study. Commented Matlab code for all analysis carried out here is available on request.

## 2.4 Materials and methods

### 2.4.1 Absolute classical conditioning

The exact experimental procedure used during conditioning of the proboscis extension response has been described elsewhere (Menzel, 2001; Stollhoff et al., 2005; Felsenberg et al., 2011). During absolute classical conditioning (data sets 1-15), honeybees were exposed to $m$ forward pairings of the conditioned stimulus (CS, odor) with the unconditioned stimulus (US, sucrose). The CS and US durations, as well as the CS-US overlaps slightly differed in the data sets (data sets 1-13: 5 sec CS, 4 sec US, 2 sec CS-US overlap; data sets 14-15: 4 sec CS, 3 sec US, 1 sec CS-US overlap). Since the CS onset preceded the US onset by a few seconds, the occurrence of the proboscis extension during this time span was documented in a binary form as the CR. For a certain trial $t$, we denote the presence (absence) of the CR with $x_t = 1 (x_t = 0)$. After conditioning, memory retention was tested by exposing the bees to the CS alone at time $T$. The animals from data sets 9-13 were tested two times at $T = 1h$ and at $T = 24h$. In this study we only included animals that (1) did not respond to the first CS during acquisition, (2) survived the entire experiment, and (3) showed the proboscis extension response elicited by sucrose feeding at the very end of the experiment. We made one exception to this rule: data sets 14 and 15 comprise 63 and 64 animals that were conditioned, but only 29 and 23 animals were tested at $T = 94h$.

### 2.4.2 Differential conditioning

In the differential conditioning paradigm (data sets 16 and 17) honeybees experienced two different odors during conditioning, the first one being rewarded (CS+) and the second one being unrewarded (CS-). Each data set comprised 12 conditioning trials (6 CS+ and 6 CS- in alternating order, starting with CS+ , ITI = 14 min), as well as a retention test for both odors at $T = 1h$ and $T = 24h$. In the data set 16, 1-hexanal and 1-octanol were used as CS+ and CS-, respectively, while in the data set 17, the odor- reward contingencies were reversed. The CS+ and CS- durations were 5 sec, the US duration was 4 sec, and the overlap between the CS+ and the US was 2 sec.

### 2.4.3 Standard data analysis

Each data set was independently analyzed by the following standard procedure: In a given trial $t$ we distinguished between group CR probabilities and subgroup CR probabilities. To compute the group CR probability $P(x_t = 1)$ we divided the number of bees that showed a CR in trial $t$ by the total number of bees $N$. Plotting $P(x_t = 1)$ against trial order $t$ results in the so-called learning curve (see, e.g., Figs. 2.1 A, 2.3 A). The subgroup probability $P(x_t = 1|x_{t-1} = 1)$ is conditioned on the expression of the CR in trial $t - 1$. It was computed by dividing the number of bees that showed a CR in both trials $t$ and $t - 1$ by the number

of bees that showed a CR in trial $t-1$. The subgroup probability $P(x_t = 1|x_{t-1} = 0)$ was computed by dividing the number of bees that showed a CR in trial $t$ but not in trial $t-1$ by the number of bees that did not show a CR in trial $t-1$. The difference between the two subgroup CR probabilities was tested for statistical significance by a $\chi^2$ test. The significance level was set to 0.05.

We specify three amendments to this analysis: (1) For the experiments in which animals were presented to two retention tests (data sets 9-13, 16, 17), the subgroup CR probabilities in the tests were always conditioned on the response in the last conditioning trial ($t = m$). (2) For the data sets 16 and 17 (differential conditioning), the analysis was separately used for CS+ and CS- trials. We introduced an apostrophe in the respective terms to indicate CS- trials. (3) For the data from the extinction experiment (a subset of 217 animals from data set 1) the analysis was used for trials $t = 1 - 8$ without any changes. The subgroup probabilities in the five retention tests were conditioned on the respective behavior in trial $t = 8$.

### 2.4.4 The simple learning-curve model

The simple learning-curve model (LCM1) hypothesizes that the gradual rise of the CR probability $P(x_t = 1)$ observed at the population level reflects the gradual rise of associative strength in individuals (see Fig. 2.4 A). Furthermore, all animals in a given population of identically treated animals are assumed to be identical with respect to learning and behavioral performance probabilities. The gradual increase of associative strength (AS) across trials is defined by the learning-rate $\epsilon$ and the asymptotic value $r$. For each data set the two parameters $\epsilon$ and $r$ are computed by fitting the equation

$$AS(t) = r(1 - \exp(-\epsilon(t-1))) \tag{2.1}$$

to the group CR probabilities $P(x_t = 1)$ on the range $t = 1...m+1$ by nonlinear regression. For the data sets 9-13, in which animals were tested twice, the trial $t = m+1$ equaled the first test, while the second test was discarded. (It should be noted that for estimating the parameters of the other two models [see below] we also did not make use of the second test.) Finally, the expression of behavior in each animal is assumed to be probabilistic, with probability $P^{LCM1}(x_t = 1) = AS(t)$ for expressing a CR, and probability $P^{LCM1}(x_t = 1) = 1 - AS(t)$ for not expressing a CR, respectively. It should be noted that the equation for the associative strength used here has an identical outcome as the Rescorla-Wagner delta rule for associative learning (Rescorla and Wagner, 1972).

### 2.4.5 The extended learning-curve model

The extended learning-curve model (LCM2) extends the simple learning-curve model by the following two features (see Fig. 2.4 B): First, learning rates can differ within a population

of identically treated animals, which is described by a discrete probability distribution $P(\epsilon_i)$. Second, for a given animal with learning-rate $\epsilon_i$ the probability for expressing or not expressing a CR at a given trial $t$ is defined by a simple threshold function:

$$P^{LCM2}(x_t = 1) = \begin{cases} 0 & \text{for } AS(t) < 0.5 \\ K & \text{for } AS(t) \geq 0.5 \end{cases} \quad \text{and} \qquad P^{LCM2}(x_t = 0) = 1 - P^{LCM2}(x_t = 1), \tag{2.2}$$

with

$$AS(t) = r(1 - \exp(-\epsilon_i(t-1))) \tag{2.3}$$

Setting the threshold to 0.5 and the asymptotic value $r$ to unity ensures an optimal dynamic range for any possible distribution of learning rates. According to the extended learning-curve model, different animals can have different associative strengths at a given trial $t$; however, the behavioral performance rule (Eq. 2.2) is assumed to be the same for all animals: Animals with an associative $AS(t) < 0.5$ will not show a CR at trial $t$, while animals with $AS(t) \geq 0.5$ will show a CR with probability $K$. The parameter $K$ is a heuristic estimate and equals the mean probability for making the observation $(x_t = 1|x_{t-1} = 1)$ across all possible trials and animals in a data set. Given these model specifications, it follows that the probability distribution of learning rates has to be estimated on the discrete learning-rate intervals: $[-\ln(1-0.5)/L, -\ln(1-0.5)/(L-1)[$, $[-\ln(1-0.5)/(L-1), -\ln(1-0.5)/(L-2)[, ...,$ $[-\ln(1-0.5)/2, -\ln(1-0.5)/1[$, $[-\ln(1-0.5)/1, \infty[$, where $L$ is the maximum number of trials in a data set. (For example, for a data set with $m$ conditioning trials and one test trial $L$ equals $m+1$.) Any two animals with learning rates from the same interval cannot be dissociated because they will have equal probabilities for any binary behavioral sequence of length $L$. The highest learning-rate interval contains all possible learning rates that would result in an associative strength equal to or larger than the threshold of 0.5 in trial 2. The second highest learning-rate interval contains all possible learning rates that would result in an associative strength equal to or larger than the threshold in trial 3, and smaller than the threshold in trial 2, and so forth. The lowest learning-rate interval contains all possible learning rates that would result in an associative strength equal to or larger than threshold in trial $L+1$ and smaller than the threshold in trial $L$. (The lower interval limit of the lowest interval was set to $-\ln(1-0.5)/L$ instead of zero in order to make the minimal resolution explicit.)

For estimating $p(\epsilon_i)$ we use a simple search and count algorithm. For each behavioral sequence in a given data set we search for the most likely learning-rate interval. The probability distribution $p(\epsilon_i)$ then equals the histogram of most likely counts for each interval, normalized by the total number of animals in a data set.

### 2.4.6 The two-state hidden Markov model

The two-state hidden Markov model hypothesizes that at each trial $t$ an animal can occupy one of two possible hidden states, $k = 1$ or $k = 2$ (see Bishop 2006 for a textbook account of hidden Markov models). Each of the two states is characterized by two probabilities for showing ($x_t = 1$) or not showing ($x_t = 0$) a CR, as well as by two probabilities for remaining in the same state or making a transition to the respective other state (see Fig. 2.4 C). In total, the model comprises 10 parameters: two a priori state probabilities $P(k_1)$, four transition probabilities between states $P(k_t|k_{t-1})$, and four observation probabilities $P(x_t|k_t)$ for observing the behavioral outcome $x_t$ stemming from state $k_t$. Five of these 10 parameters are independent and were estimated by the Baum-Welch algorithm (Welch 2003) implemented in the hidden Markov model toolbox by Kevin Murphy for Matlab (University of British Columbia, Canada). Parameter estimation was insensitive to the choice of initial parameters for the data sets 1-15.

### 2.4.7 Cross-validation

For each data set we computed the average log likelihood of the three models after 50 rounds of cross-validation. At each round, the data was split into four sets of equal size, each of which was then used once for testing and three times for training.
Data analysis and modeling was performed in Matlab (The MathWorks).

## 2.5 Tables and table captions

| Data set | $m$ | ITI (min) | $T(h)$ | $N$ | Odor | Data origin |
|---|---|---|---|---|---|---|
| 1 | 3 | 10 | 24 | 517 | Carnation oil | Stollhoff et al. (2005) |
| 2 | 3 | 10 | 25 | 98 | Carnation oil | Stollhoff et al. (2005) |
| 3 | 3 | 10 | 26 | 113 | Carnation oil | Stollhoff et al. (2005) |
| 4 | 3 | 10 | 28 | 92 | Carnation oil | Stollhoff et al. (2005) |
| 5 | 3 | 10 | 48 | 85 | Carnation oil | Stollhoff et al. (2005) |
| 6 | 3 | 10 | 72 | 94 | Carnation oil | Stollhoff et al. (2005) |
| 7 | 3 | 10 | 24 | 87 | Carnation oil | KB Gehring and D Eisenhardt,unpubl. |
| 8 | 3 | 2 | 24 | 58 | Carnation oil | L Morgenstern, J Felsenberg, and D Eisenhardt, unpubl. |
| 9 | 3 | 2 | 24, 48 | 64 | Carnation oil | V Antemann, J Felsenberg, and D Eisenhardt, unpubl. |
| 10 | 4 | 30 | 1, 24 | 56 | Isoamyl Acetate | NK Chakroborty,unpubl. |
| 11 | 4 | 30 | 1, 24 | 66 | Isoamyl Acetate | NK Chakroborty,unpubl. |
| 12 | 5 | 30 | 1, 24 | 48 | 7-pentadecene | NK Chakroborty,unpubl. |
| 13 | 5 | 30 | 1, 24 | 37 | 6-pentadecene | NK Chakroborty,unpubl. |
| 14 | 12 | 0,5 | 96 | 63 | Hexanol | Menzel (2001) |
| 15 | 12 | 15 | 96 | 64 | Hexanol | Menzel (2001) |
| 16 | 6,6' | 15 | 1,24 | 48 | Hexanol, octanol' | NK Chakroborty,unpubl. |
| 17 | 6,6' | 15 | 1,24 | 50 | Octanol, hexanol' | NK Chakroborty,unpubl. |

**Table 2.1** Data analyzed from absolute (data sets 1-15) and differential (data sets 16-17) conditioning experiments. The 17 data sets differed with respect to several experimental parameters. ($m$) Number of conditioning trials in the acquisition session, ITI inter-trial interval in the acquisition session; ($T$) time of the retention test; ($N$) number of animals in each data set. For the data sets 16 and 17, the apostrophe indicates CS- trials. The data were contributed by the individual co-investigators as indicated.

| **Data set** | $\log(P^{LCM2}) - \log(P^{LCM1})$ | $\log(P^{LCM2}) - \log(P^{HMM})$ |
|---|---|---|
| 1 | 19.3 | 0.0 |
| 2 | 3.6 | 0.4 |
| 3 | 5.4 | 1.2 |
| 4 | 3.3 | 0.7 |
| 5 | 2.7 | 0.5 |
| 6 | 5.0 | 0.9 |
| 7 | 1.7 | 0.9 |
| 8 | 4.0 | 0.8 |
| 9 | 2.3 | 0.9 |
| 10 | 6.0 | -0.2 |
| 11 | 6.3 | 0.3 |
| 12 | 12.6 | 1.9 |
| 13 | 14.6 | 3.4 |
| 14 | 14.0 | -20.9 |
| 15 | 31.4 | -17.8 |

**Table 2.2** Differences between the mean log-likelihoods of the three models after 50 rounds of fourfold cross-validation. (LCM1) Simple learning-curve model; (LCM2) extended learning-curve model; (HMM) two-state hidden Markov model. The extended learning-curve model was chosen as a reference point because it was the best model for 11 of the 15 data sets. The extended learning-curve model performs better than the simple one in all 15 cases, and it also performs better than the hidden Markov model in 11 out of 15 cases.

## 2.6 Figures and figure captions

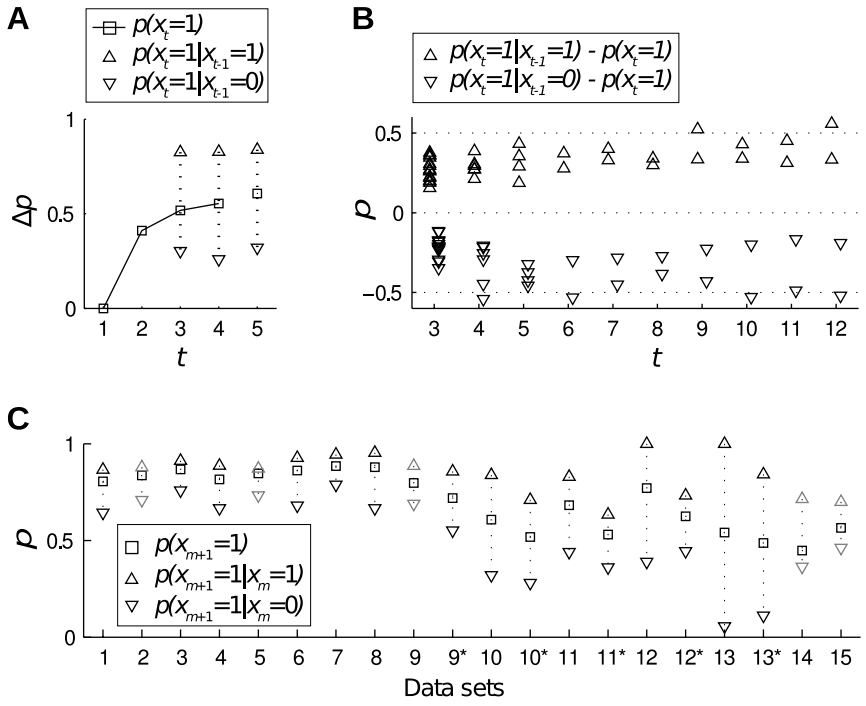


**Figure 2.1** For data sets 1-15, the group CR probability is a poor estimate for individual behavior during conditioning and in the memory retention tests. Animals that responded in the previous trial $t-1$ always showed a higher probability for responding in trial $t$ than animals that did not respond in the previous trial $t-1$. (A) Group CR probabilities $P(x_t = 1)$ and the two subgroup CR probabilities $P(x_t = 1|x_{t-1} = 1)$ and $P(x_t = 1|x_{t-1} = 0)$ for data set 11. The data were analyzed by the standard data analysis (see Materials and Methods). Black triangles indicate a significant difference between the two subgroup probabilities in the respective trial ($\alpha = 0.05$). Gray triangles indicate a nonsignificant difference. (B) Subgroup CR probabilities $P(x_t = 1|x_{t-1} = 1)$ and $P(x_t = 1|x_{t-1} = 0)$ computed for all conditioning trials and for all 15 data sets. For display, we subtracted the group CR probabilities from the subgroup CR probabilities. (C) Subgroup CR probabilities during memory retention conditioned on the outcome of the final conditioning trial computed for data sets 1-15. For data sets 9-13, the asterisk indicates the second memory retention test (see Table 2.1).

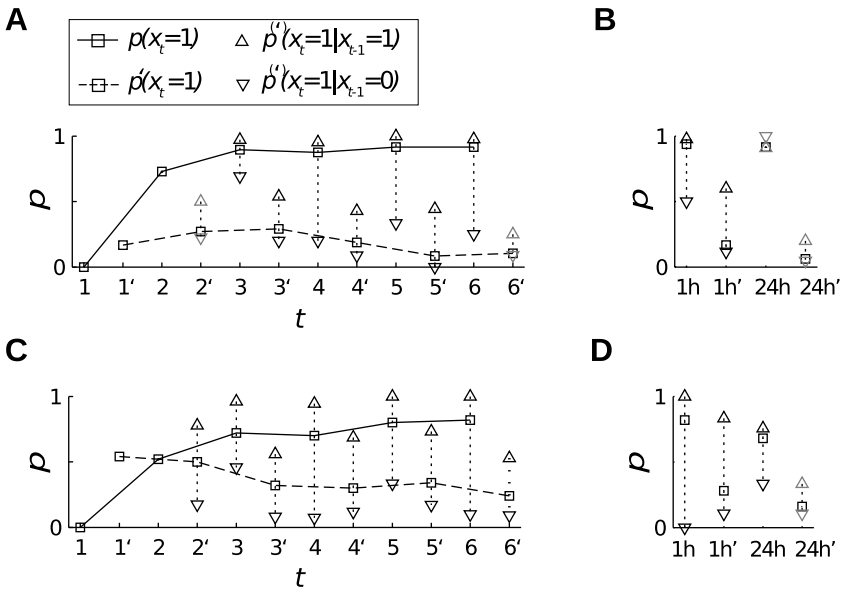


**Figure 2.2** During differential conditioning, animals that responded to the CS+ (CS-) in the previous trial $t-1$ ($t'-1$) showed a higher probability for responding to the CS+ (CS-) in trial $t$ ($t'$) than animals that did not respond to the CS+ (CS-) in the previous trial. The apostrophe indicates CS- trials. (A) Group and subgroup probabilities during conditioning for data set 16. Behavioral responses in CS+ and CS- trials were independently analyzed by the standard data analysis (see Materials and Methods). Black triangles indicate a significant difference between subgroup probabilities ($\alpha = 0.05$). Gray triangles indicate a nonsignificant difference. (B) Animals from data set 16 were repeatedly tested for memory retention by presenting the CS+ and the CS- at $T = 1h$ and $24h$. The subgroup probabilities were conditioned on the response to the last CS+ (CS-) presentation during conditioning. (C,D) Same analysis as in A and B, but for a different group of animals (data set 17), which was differentially conditioned to the reversed odor-reward contingencies.

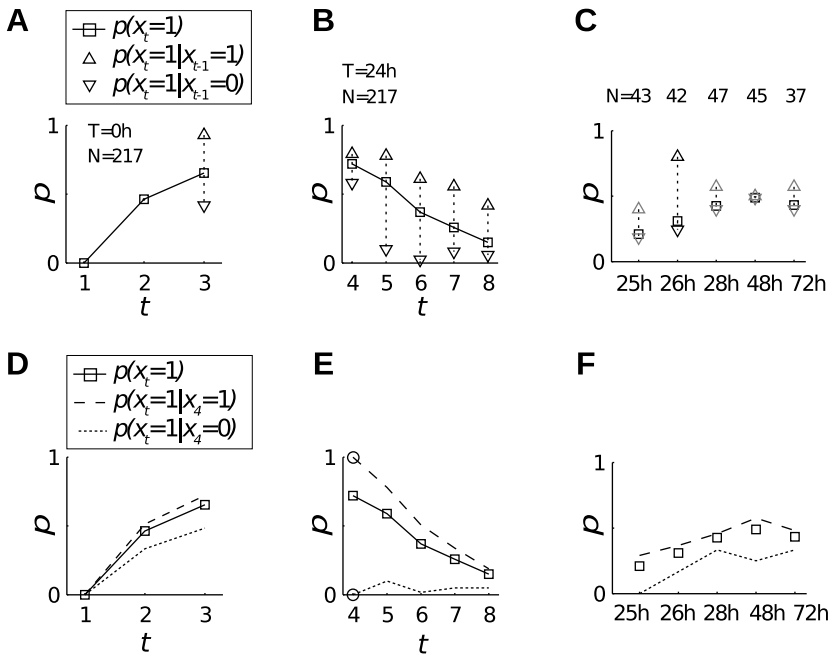


**Figure 2.3** A subset of animals ($n = 217$) from data set 1 was presented to five extinction trials 24 h after conditioning (A,B). After extinction ($t > 8$), animals were divided into five groups, each of which was tested for spontaneous recovery at a different time point (C). (A-C ) During acquisition, extinction, and the retention tests, animals that responded in the previous trial $t - 1$ showed a higher probability for responding in trial $t$ than animals that did not respond in the previous trial. The standard data analysis was used without change from trial $t = 1...8$ (see Material and Methods). The subgroup probabilities in the five retention tests were conditioned on the response in trial $t = 8$. Black triangles indicate a significant difference between subgroup probabilities ($\alpha = 0.05$). Gray triangles indicate a nonsignificant difference. (D-F) Group CR probabilities as in A, B, and C. The two dashed curves show the response probabilities for the two disjoint subgroups defined by their behavior in trial $t = 4$. The employed selection criterion (black circles) results in a steady performance difference between the two subgroups of animals in all other trials.

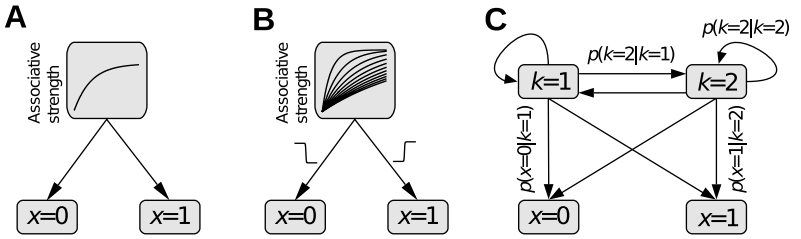


**Figure 2.4** Three model hypotheses for the observed binary sequences of conditioned responses. (A) The simple learning-curve model (LCM1) hypothesizes that the gradually increasing group CR probability $P(x_t = 1)$ reflects the gradually increasing associative strength in individual animals. All animals are the same and at a given trial t express the CR with probability $P^{LCM}(x_t = 1)$. (B) The extended learning-curve model (LCM2) assumes that individual animals in a given group can differ with respect to their learning rates, which is described by a discrete probability distribution $P(\epsilon_i)$. The probability for observing or not observing a CR is described by a simple performance function that is the same for all animals of the group. (C ) The hidden Markov model assumes that individual animals can be in one of two discrete hidden states. At the beginning of the conditioning experiment all animals are in the naive or unlearned state $k = 1$. At each conditioning trial, animals in this state have a certain success rate $P(k = 2|k = 1)$ for making a transition to the learned state $k = 2$. The probability $P(k = 2|k = 2)$ for remaining in the learned state is typically very high. The probability $P(x = 1|k = 2)$ for expressing a CR is high in the state $k = 2$, while the CR probability $P(x = 1|k = 1)$ is low in the state $k = 1$. (See Materials and Methods for a description of how parameters are estimated for the three models.)

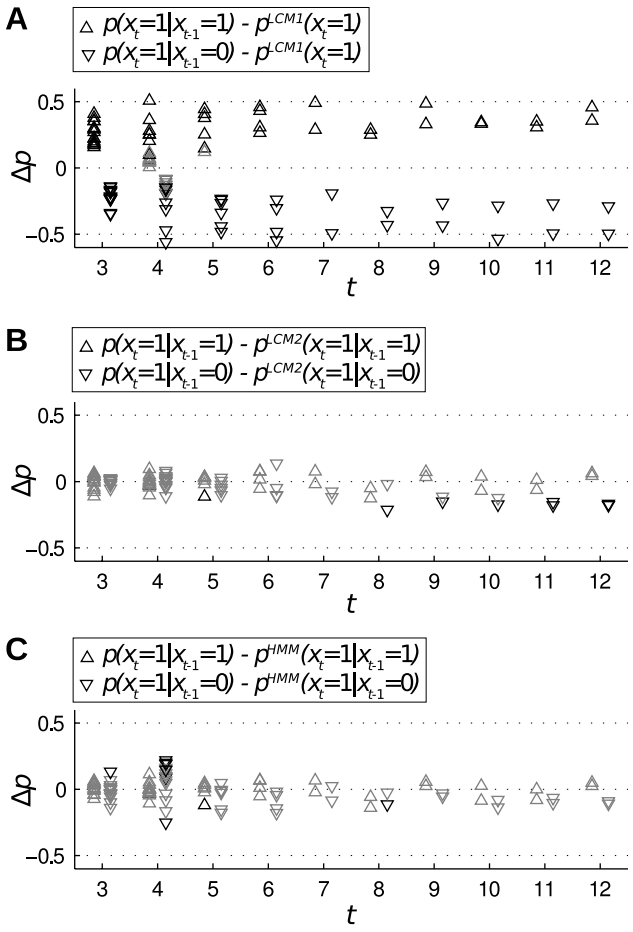


**Figure 2.5** Differences between the empirical subgroup's CR probabilities and the respective model estimations for data sets 1-15. The differences were tested for statistical significance with a two-tailed binomial test. Black triangles denote statistical significance ($\alpha = 0.05$). Gray triangles denote no statistical significance. (A) The simple learning-curve model (LCM1) cannot capture the experimental data at the level of subgroups. (B,C ) Both the extended learning-curve model (LCM2, B) and the hidden Markov model (HMM, C) provide a good fit for the experimental data.

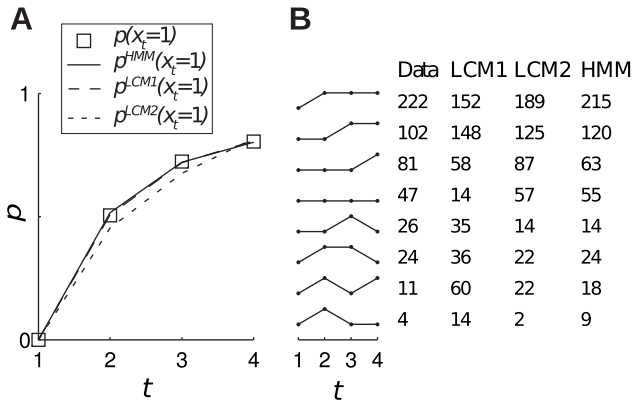


**Figure 2.6** Illustration of the performance differences between the simple learning-curve model (LCM1), the extended learning-curve model (LCM2), and the hidden Markov model (HMM) for data set 1 ($n = 517$). (A) All three models can describe the behavioral data at the population level. (B) The data set 1 contains eight different binary behavioral sequences. (Left column) Absolute sequence frequencies as counted in the data. (Middle and right column) Absolute sequence frequencies as computed by the three models rounded to integers. The distribution of sequences in the experimental data is best captured by the hidden Markov model (75 errors), followed by the extended learning-curve model (99 errors) and the simple learning-curve model (252 errors). Errors are the sum over the absolute differences between frequencies in the data and in the respective model prediction.

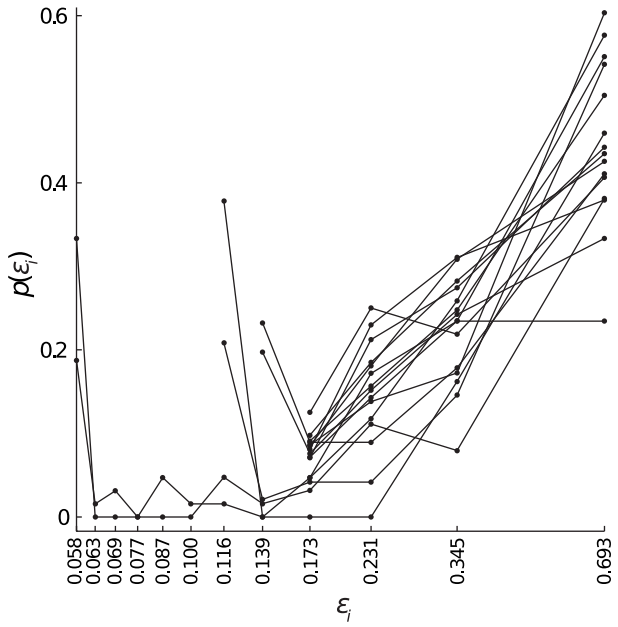


**Figure 2.7** Empirical probability distribution estimates $P(\epsilon_i)$ for data sets 1-15. The x-axis has been scaled logarithmically. The values of $\epsilon_i$ on the x-axis always denote the lower limits of the respective learning-rate intervals (see Materials and Methods, "The extended learning-curve model").

# Chapter 3

# Rapid learning dynamics in individual honeybees during classical conditioning

**Abstract**

Associative learning in insects has been studied extensively by a multitude of classical conditioning protocols. However, so far little emphasis has been put on the dynamics of learning in individuals. The honeybee is a well-established animal model for learning and memory. We here studied associative learning as expressed in individual behavior based on a large collection of data on olfactory classical conditioning (24 datsets, 3,005 animals). We show that the group-averaged learning curve and memory retention score confound three attributes of individual learning: the ability or inability to learn a given task, the generally fast acquisition of a conditioned response in learners, and the high stability of the conditioned response during consecutive training and memory retention trials. We reassessed the prevailing view that more training results in better memories and found that multiple-trial conditioning did not enhance 24h memory retention or discriminative power over single-trial conditioning in individuals. Previous studies suggested that honeybees vary in learning performance due to subjective differences in perceived reinforcement by the sucrose reward. Following this hypothesis we explain how inter-individual differences in learning can be accommodated within the Rescorla-Wagner theory of associative learning. In both data-analysis and modeling we demonstrate how the conflict between population-level and single-animal perspectives on learning and memory can be disentangled.

## 3.1 Introduction

Classical conditioning relies on the assumption that changes in conditioned response (CR) probability observed during training adequately represent neuronal plasticity (Dubnau et al., 2003; Dudai, 2004) Commonly, behavioral plasticity is quantified by averaging over a population of identically treated animals. However, average performance scores can obscure the learning dynamics in individuals (Gallistel et al., 2004). Animals in a given sample can vary considerably in several attributes of individual learning, such as in the ability to learn a task, the speed of learning, the asymptotic performance (Dukas, 2008) and the latency until a CR is initiated for the first time (Gallistel et al., 2004). To give a simple example: A group of animals may only contain individuals exhibiting optimal performance as well as individuals lacking the ability to learn a given task. Averaging will then misleadingly display a moderate learning performance. To give a second example: Averaging over individuals that exhibit rapid changes in response probability at different time-points during training will yield a gradual learning curve at the population level, leading to the false assumption of a gradual performance improvement over training trials in individuals (Gallistel et al., 2004).

Based on a large collection of classical conditioning data (Table 3.1) we here studied the disparity between individual and group-average learning dynamics in the honeybee (*Apis mellifera*). Classical conditioning of the proboscis extension response (PER) combined with electrophysiology, biochemistry and pharmacology has proven to be a powerful tool for studying the neuronal mechanisms of learning and memory in this animal model (Menzel, 2001; Schwärzel and Müller, 2006; Giurfa, 2007; Giurfa and Sandoz, 2012; Menzel, 2012). Our study follows four objectives: (1) Previous analysis has shown that average behavior does not represent individual behavior in the honeybee (Pamir et al., 2011). Alternative to the prevailing use of group-average learning curves and retention scores we present a novel parametric description for learning in individuals. (2) The current model of memory phases in the honeybee distinguishes between single-trial and three-trial induced memories (Menzel, 1999; Müller, 2012). Using a novel experimental protocol we tested the hypothesis that differences in memory retention after single-trial and three-trial conditioning at the group-average level are caused by a recruitment effect, that is three-trial conditioning may allow more animals to acquire a conditioned response than single-trial conditioning. (3) We analyzed the modulation of individual learning parameters under altered training conditions, such as in differential conditioning, delay conditioning and massed conditioning trials. (4) Within-group heterogeneity in learning is often neglected in theoretical models. For the honeybee, several studies showed that factors such as satiation level, behavioral role or age have an effect on individual sensitivity for sucrose, which in turn affects learning performance (Friedrich et al., 2004; Scheiner et al., 1999; Behrends and Scheiner, 2012). We implemented the hypothesis that honeybees vary in learning performance due to a differential evaluation of the sucrose reward (Scheiner et al., 2005) within the Rescorla-Wagner model for associative

learning (Rescorla and Wagner, 1972).

## 3.2 Materials and methods

### 3.2.1 Classical conditioning of the proboscis extension response in the honeybee

Olfactory classical conditioning of the proboscis extension response (PER) in the honeybee has been described in detail (Bitterman et al., 1983; Scheiner et al., 2003; Stollhoff et al., 2005; Szyszka et al., 2011; Felsenberg et al., 2011). Briefly, during the conditioning session, a group of $n$ animals was individually exposed to $m$ forward pairings of the conditioned stimulus (CS, odor) with the unconditioned stimulus (US, sucrose). Memory retention was measured by presenting the CS alone at time point $T(h)$ after conditioning. Only animals responding to sucrose alone at the end of the experiment were included in the analysis. Typically the CS duration was in the range of 3-5 seconds, the US duration equaled 3-5 seconds, and the CS-US overlap equaled 1-2 seconds. The occurrence of the proboscis extension during the CS not overlapping with the US was documented in a binary form as the conditioned response (CR). For conditioning trial $t$ we denote the absence or presence of the CR with $x_t = 0$ or $x_t = 1$ respectively. Table 3.1 provides an overview over the experimental data analyzed in this study. Details for each dataset are provided in the following.

### 3.2.2 Experimental data

#### Absolute conditioning (datasets 1-12)

Datasets 1-12 comprise data on olfactory classical conditioning. Animals in data sets 1, 10, 11 and 12 were tested twice for memory retention (see Table 3.1). For consistency we do not further regard the first test. CS duration, US duration and CS-US overlap equaled 5, 4 and 2 seconds respectively.

#### Trace and delay conditioning (dataset 13-15)

Dataset 13 comprises data on trace conditioning (compare with Figure 2Aii (trace) in Szyszka et al., 2011). CS duration and US duration equaled 0.5 and 3 seconds respectively. The CS and the US did not overlap. The gap between CS offset and US onset was 4.5 seconds. Dataset 14 comprises data on delay conditioning (compare with Figure 2Aii (delay) in Szyszka et al., 2011). CS duration, US duration and CS-US overlap equaled 6, 3 and 1 seconds respectively. Dataset 15 comprises data in which the time difference between the onset of the CS and the US was systematically varied in 8 subgroups of animals (compare with Figure 2Bii in Szyszka et al., 2011, CS-US onset differences equaled -6, 0, 1, 2, 3, 6, 10 and 15 seconds). CS durations and US durations equaled 0.5 and 3 seconds respectively.

### Olfactory and tactile conditioning (dataset 16, 17)

Dataset 16 and 17 comprise data on olfactory and tactile conditioning (compare with Table 1 in Scheiner et al., 2001a). Here we did not differentiate between high-strain and low-strain bees. For tactile conditioning small rectangular copper plates with vertical grooves were used as the CS (for details see Erber et al., 1998; Scheiner et al., 1999, 2001a) and sucrose was used as US and reward. The US was the same in olfactory and tactile conditioning. Prior to the conditioning session individuals were tested for their responsiveness to sucrose by touching their antennae with 9 different sucrose concentrations (1, 1.6, 2.5, 4, 6.3, 10, 16, 25 and 40% (w/v)). Between the sucrose stimulations, antennae were touched with water to test for sensitization effects. The inter-trial-interval was 2 minutes to avoid intrinsic sensitization. For each animal the total number of proboscis responses to the first water and the nine sucrose stimulations was counted. This sum is referred to as the gustatory response score (GRS) of a bee (Scheiner et al., 2004). In the conditioning session, animals were trained by 10 pairings of CS (citral, 2 $\mu l$ added to airstream for 3 seconds before onset of the sucrose stimulation) and US (0.2 $\mu l$ 30% sucrose solution) at an inter-trial-interval of 5 minutes. 24 hours after conditioning, bees were exposed to five unreinforced CS. In the present analysis we only included the first CS only trial as a memory retention test and disregarded all subsequent trials.

In each trial, the CS was given 3 seconds before the onset of the US at the antennae, which was followed by a proboscis stimulation with sucrose. The CS-US overlap was 1s and the US duration at the proboscis was 1s. It should be noted that for dataset 16 and 17 equal proportions of animals from different ranges of gustatory response scores were collected. Consequently, neither of the two datasets is treated as an unbiased sample of animals in our analysis.

### Massed and spaced conditioning (datasets 18, 19)

Datasets 18 and 19 comprise animals from massed and spaced training conditions (Menzel et al., 2001). Under massed training conditions inter-trial-intervals equaled 30 seconds, while under spaced training conditions inter-trial-intervals equaled 15 minutes. We included all animals that survived the conditioning session in our analysis (group sizes differ from Menzel et al., 2001). CS duration, US duration and CS-US overlap equaled 4, 3 and 1 seconds, respectively.

### Differential conditioning (datasets 20)

Dataset 20 comprises data on differential classical conditioning where two groups of animals were conditioned by 6 rewarded (CS+) and 6 unrewarded (CS-) odor presentations. The first group received 1-hexanal and 1-octanol as CS+ and CS- respectively, while in the second

group the odor reward contingencies were reversed. Conditioning started with a CS+ trial and then alternated between CS- and CS+. The inter-trial-interval between identical stimuli equaled 14 minutes. Animals were tested for memory retention and discrimination at 1 and 24 hours. CS+ (CS-) duration, US duration and CS-US overlap equaled 5, 4 and 2 seconds respectively.

### 3.2.3 Experiment 1 (datasets 21, 22)

Experiment 1 was performed in the summer of 2011 with honeybee foragers (*Apis mellifera*) from outside hives. During classical conditioning bees were trained to associate either 1-hexanol or 1-nonanol (CS) with a 1 M sucrose reward (US). The odorants were diluted 1:100 in mineral oil (Sigma-Aldrich, Deisenhofen, Germany), and were presented as 4-second long stimuli with a custom-made olfactometer (Szyszka et al., 2011). US duration and CS-US overlap equaled 3 and 2 seconds respectively.

Experiment 1 was designed to obtain bees that fall into one of the following 4 subgroups: 0111, 01, 001, 0001. The binary notation equals the sequence of conditioned responses during the conditioning session, referred to as the CR-history. The leftmost number equals the CR in the first conditioning trial, while the rightmost number equals the CR in the last conditioning trial. To obtain these subgroups at comparable sample sizes we chose the following experimental protocol: In each experimental run, 16 bees were conditioned in parallel. Out of these bees, four animals were conditioned four times without interfering with the conditioning process (dataset 21). Another eight bees were conditioned until the first CS-evoked proboscis extension, yielding the CR histories 01, 001, 0001 and 0000. The remaining 4 bees were conditioned 4 times in case they showed a proboscis extension in the second trial or until the first proboscis extension otherwise. The 12 animals per plate conditioned by the latter two protocols are referred to as dataset 22.

Memory retention was tested 24 hours after training. During the test, each bee was stimulated with the CS and a new odorant which in addition allowed the calculation of a discrimination index ($DI$, see below) by subtracting the response to a new odorant from the response to the CS (Biergans et al., 2012; Matsumoto et al., 2012). This procedure eliminates all non-associative effects of the conditioning procedure, such as sensitization or pseudoconditioning, which would also increase animals responsiveness (Tully, 1984). 1-hexanol and 1-nonanol were equally often used as CS and new odorant. For each behavioral response we also recorded its duration to capture possible differences in memory strength (Smith and Menzel, 1989a). Response duration was measured as time (in one-second intervals) between the beginning of the proboscis extension until its first retraction below the imagined horizontal line. In case of no response, no duration value was incorporated. The inter-trial interval was 10 minutes both in training and in the test.

The discrimination index $DI$ was computed as

$$DI = \frac{1}{N}\sum_{i=1}^{N} x^i_{CS} - x^i_{new} \tag{3.1}$$

where $x^i$ denotes the CR of animal $i$ to the presentation of the CS and new odorant, and $N$ equals the number of animals in a given subgroup defined by the CR history. The discrimination index based on the CR duration was computed as

$$DI_{dur} = \frac{1}{N}\sum_{i=1}^{N} \frac{d^i_{CS} - d^i_{new}}{max(d^i_{CS}, d^i_{new})} \tag{3.2}$$

where $d^i$ denotes the duration of the proboscis extension of animal $i$ to the CS and new odorant. Differences between durations of proboscis extensions were normalized individually by the maximum duration of animal $i$ to either stimuli.

### 3.2.4 Experiment 2 (datasets 23, 24)

Experiment 2 was performed in late autumn/winter 2011 with honeybee foragers (*Apis mellifera*) from indoor hives. Animals either experienced two-trial conditioning (dataset 23) or single-trial conditioning plus a CS presentation without sugar reward 10 minutes after conditioning (dataset 24). This yielded the CR-history subgroups 01 and 0(1). The bracket notation indicates the CR in the CS-only trial. Memory retention and discriminatory power was measured as described in Experiment 1, and 1-hexanol and 1-nonanol were used equally often as CS and new odorant. Animals of data sets 23 and 24 were conditioned in parallel.

### 3.2.5 Data analysis

An example raw dataset of binary CRs from absolute conditioning is depicted in Figure 3.1 A. The data was analyzed by the following standard procedure: The notation $x_t$=1 ($x_t$=0) denotes the presence (absence) of the CR on trial $t$. The trial index $t$ ranges from 1 to the maximum number of trials, including the memory test. The average CR probability equals the percentage of animals showing a CR in trial $t$. Average CR probabilities across trials were fitted by the equation

$$p(CR) = a(1 - e^{-b(t-1)}) + c(t-1) \tag{3.3}$$

where the three free parameters $a$, $b$ and $c$ were estimated by least-squares minimization. The point in trial time at which the regression curve assumes its maximum $p_{max}$ is denoted as $t_{max}$ (Figure 3.1 B). Animals in each dataset were divided into disjunctive subgroups defined by the trial $t_{firstCR}$ at which animals showed their first CR (see Figure 3.1 C for a histogram of first CRs). Animals that did not show a response in any of the trials constituted

the subgroup of non-responders. For all occurring first CR indexes $j$, we computed the conditional probabilities $p(x_t{=}1|t_{firstCR}{=}j)$ with $t > j$. Figure 3.1 B exemplifies this analysis for $p(x_t = 1|t_{firstCR} = 2)$ and $p(x_t = 1|t_{firstCR} = 3)$. Taking the mean over all conditional probabilities $p(x_t{=}1|t_{firstCR}{=}j)$ with $t > j$ results in the CR stability of a subgroup defined by $t_{firstCR} = j$. Taking the weighted mean of the CR stabilities of all subgroups results in the overall CR stability of a given dataset. The CR stabilities of subgroups were weighted according to subgroup sizes. The CR stability is a measure of how constantly individuals of a given dataset responded once they had started to respond. From the definition follows that neither bees that do not show a CR in any of the trials (non-responders), nor animals that only respond in the last trial contribute to this parameter.

### 3.2.6 The classical Rescorla-Wagner model

The Rescorla-Wagner (RW) model assumes that associative learning during classical conditioning is driven by prediction errors (Rescorla and Wagner, 1972; Sutton and Barto, 1990). At each conditioning trial $t$ the animal experiences a prediction error $(\lambda - v_t)$ defined as the difference between the maximum associative strength $\lambda$ supported by the unconditioned stimulus (US), and the associative strength $v_t$ of the conditioned stimulus (CS) at the current trial. In the following we refer to the parameter $\lambda \in [0, 1]$ as the US effectiveness. After each trial, the associative strength $v_t$ is updated according to the rule

$$v_{t+1} = v_t + \alpha(\lambda - v_t) \tag{3.4}$$

where $\alpha \in [0, 1]$ is the learning rate, defined in the original theory as the product of CS and US salience (Rescorla and Wagner, 1972). The update rule leads to a gradual strengthening of associative strength across conditioning trials (3.6 A). Here we assume a linear mapping between associative strength and CR probability, hence the probability of animal $i$ to show a CR on trial $t$ is $p^i(x_t{=}1) = v_t$, and the probability for not showing a CR is $p^i(x_t{=}0) = 1 - v_t$. In Equation 3.4 the value $v_t$ denotes the associative strength at precisely the time of trial $t$, hence before the actual learning induced in this trial has become effective. This is analogous to the experimental situation in which the behavior observed in trial $t$ is taken as a monitor of the associate strength induced in all previous trials.

The two free parameters $\alpha$ and $\lambda$ were estimated by minimizing the negative log-likelihood of the model on a given dataset by the L-BFGS-B algorithm for bound constrained optimization (Byrd et al., 1994; Zhu et al., 1997). The bounds for $\alpha$ and $\lambda$ were set to $[0, 1]$. The starting value for $\alpha$ and $\lambda$ were determined by a grid search on the range $[0, 1]$ with a grid distance of 0.1.

### 3.2.7 The Rescorla-Wagner model with heterogeneous learning performance

In order to account for heterogeneous learning performance within a group of identically treated animals we considered two simple extensions of the classical Rescorla-Wagner model: (1) Different animals learn at equal rates, but vary in their experience of the same physical US. (2) Different animals experience a common US effectiveness, but vary in their learning rates. For the first case we described the heterogeneity in US effectiveness $\lambda$ by estimating a probability distribution $P(\lambda)$: For each binary behavioral sequence in a given dataset, we estimated the probability density function of this sequence over the range $\lambda \in [0, 1]$. The normalized sum of these individual probability density functions then equals the probability distribution $P(\lambda)$. The most likely learning rate $\alpha \in [0, 1]$ for a given dataset was estimated by minimizing the negative log-likelihood of the model by the L-BFGS-B algorithm for bound constrained optimization (Byrd et al., 1994; Zhu et al., 1997). The bounds for $\alpha$ were set to $[0, 1]$ and the starting value for $\alpha$ was determined by a grid search on the range $[0, 1]$ with a grid distance of 0.1.

For the second model (heterogeneity in learning rate) the probability distribution $P(\alpha)$ and the common US effectiveness $\lambda$ were estimated fully analogous to the first model. We computed the eligibility of the classical Rescorla-Wagner model and of the two extended models by a four-fold cross-validation algorithm. Data analyses were carried out in Python.

## 3.3 Results

### 3.3.1 The conditioned response is stable within individual honeybees

A typical dataset of binary behavioral responses (CR matrix) from absolute classical conditioning is depicted in Figure 3.1 A, while Figure 3.1 B and C exemplify the performed data analysis. We first asked how persistently individuals kept responding during conditioning once they had shown their first response. Quantifying this behavioral feature by a parameter termed CR stability (see Materials and Methods), we found that the mean CR stability across all datasets with standard training conditions equaled $(86.4 \pm 6.5)\%$ (datasets 1-12, 14, 19, 21). Hence once individuals had elicited their first response, they kept responding in all subsequent trials with a high probability (Figure 3.1, 3.2; Table 3.2).

Dissecting the CR stability further we found a trend that animals responding early had a higher CR stability than animals responding later: The CR stability of animals showing their first CR on the second trial equaled $(89.7 \pm 5.4)\%$ (datasets 1-12, 14, 19, 21), for animals showing their first CR on the third trial it equaled $(83.4 \pm 10.3)\%$ (datasets 1-12, 14, 19, 21), and for animals showing their first CR on the forth trial it equaled $(66.3 \pm 14.2)\%$ (datasets 10, 12, 14, 19, 21). In individual datasets this overall decrease was seen in 11 out of 15 cases between the second and third trial, and in four out of five cases between the third and fourth trial.

### 3.3.2 54% of the animals start to respond after a single conditioning trial

Next we analyzed at which trial individuals typically showed their first CR. Histograms of first CRs in trial time are displayed in Figure 3.1 C and in the lower panels of Figure 3.2. We found that $(54.1 \pm 11.4)\%$ of the animals which showed at least one response in any of the trials started to respond in the second trial, i.e. after having experienced a single CS-US paring. (datasets 1-12, 14, 19, 21). By the third trial $(80.6 \pm 8.7)\%$ of all responding animals had started to respond, and by the forth trial $(95.9 \pm 6.4)\%$ had started to respond. On average, the first CR was shown after $2.8 \pm 0.4$ trials (datasets 1-12, 14, 19, 21, Table 3.2). First CR histograms of datasets with many conditioning trials (5 to 12) furthermore imply that there is a population of animals that cannot be recruited even by prolonged conditioning (Figure 3.2 E-L, black histogram bars denote non-responders). The average percentage of non-responders equaled $(21.5 \pm 7.6)\%$ (datasets 10, 12, 14, 19, 21, only datasets in which the maximum of the regression curve (Equation 3.3) was reached during conditioning were considered).

In some of the raw datasets animals that extended their probosces to the first CS presentations were not excluded by the experimenter (total $N$=96, datasets 3-12, 14). For consistency we have so far not included these animals in our analysis. We asked if these spontaneous responders ($t_{firstCR}$=1) reliably responded to the CS in subsequent trials. We found that this was the case by computing the CR stability of these animals, which equaled $(85.0 \pm 12.8)\%$.

We furthermore found that spontaneous responders discriminated well between the CS+ and the CS- during differential conditioning as well as in a subsequent memory test (21 animals of dataset 20, Figure 3.3).

### 3.3.3 The average learning curve reflects the recruitment of individual honeybees with a stable CR

As described in the previous two sections, individual animals acquired a stable conditioned response within the first few conditioning trials. However, a portion of animals did not respond in any of the trials. How can these learning dynamics in individuals be reconciled with the learning dynamics apparent at the population level, often referred to as the "learning curve" or "acquisition function"? We described the learning dynamics at the population level by two parameters: The maximum $p_{max}$ of the regression curve (Equation 3.3) on the average CR probabilities, typically referred to as the asymptote of learning, and the position of this maximum $t_{max}$ in trial time (Materials and Methods, this analysis is exemplified in Figure 3.1 B). We found that $t_{max}$ reflected a saturation in the recruitment of honeybees showing a stable CR. By this time-point, $(91.3 \pm 5.2)\%$ of the animals that showed at least one CR in any of the trials had started to respond (datasets 10, 12, 14, 19, 21, only datasets in which the maximum was reached during conditioning were considered).

The value of $p_{max}$ reflected two behavioral characteristics at the level of individuals, the proportion of non-responding animals ($N_{non-responders}/N$) and the CR stability. We found that the following rule of thumb

$$p_{max} \approx CR_{stability}\,(1 - N_{non-responders}/N). \tag{3.5}$$

holds for all datasets with only a small error of $(0.8 \pm 1.8)\%$ (datasets 10-12, 14, 16, 19, 21). This relation illustrates that the parameter $p_{max}$ does not equal a performance asymptote of individual learning, instead it represents the percentage of animals having acquired a CR during conditioning, modulated by the $CR_{stability}$ of these animals.

The same finding applies to the memory retention test. The group-average CR probability in the retention test did not represent memory retention in individual honeybees. For animals that showed at least one CR in any of the conditioning trials the CR probability in the retention test equaled $(72.0 \pm 6.7)\%$ (datasets 10, 11, 12, 14, 19*, 21. Only datasets in which the maximum $p_{max}$ was assumed during conditioning were taken into account. Dataset 19* consisted of a subgroup of animals from dataset 19 that survived until the retention test at 72h). However, memory retention in animals that did never respond during conditioning equaled only $(24.2 \pm 13.7)\%$.

### 3.3.4 24h memory retention and discriminatory power did not differ for multiple-trial and single-trial conditioning

How does the number of conditioning trials affect the stability or strength of the induced memory? The prevailing hypothesis states that three-trial conditioning induces a stable memory, expressed in a high CR probability 24h after training, whereas single-trial conditioning induces a weaker memory with a low 24h retention probability (Menzel, 1990, Fig. 9.8; Müller, 2012, Fig. 1; Menzel, 2012, Fig. 2).We asked whether the commonly found difference between group-average retention probability after three-trial and single-trial conditioning indeed reflects enhanced memory retention in individuals after more training, or whether it actually reflects a recruitment effect. The term recruitment effect refers to the hypotheses that three-trial conditioning may allow more animals to acquire a stable CR during conditioning, whereas single-trial conditioning may allow fewer animals to acquire a stable CR.

In order to study the effect of single-trial and multiple-trial conditioning on 24h memory retention and discriminatory power in individuals we carried out two experiments (Experiments 1 and 2, see Materials and Methods). We found that memory retention after four-trial conditioning in individuals with a CR-history of 0111 did not significantly differ from memory retention after two-trial conditioning (CR-history 01) (Figure 3.4 Ai). The CR-history denotes the sequence of CRs during conditioning with the symbols 0 (no response) and 1 (response). The leftmost symbol represents the outcome of the first conditioning trial and the rightmost symbol represents the outcome of the last conditioning trial. In addition we found that memory retention after two-trial conditioning (CR-history 01) did not significantly differ from memory retention after single-trial conditioning (CR-history 0(1)) (Figure 3.4 Bi, Bii). A CR-history of 0(1) denotes animals that experienced one CS-US pairing in the first trial, and extended their proboscis to an unrewarded CS in the second trial of the training phase.

We obtained the same result when looking at the discriminatory power of the induced 24h memories (Figure 3.4 Bii): The discrimination index (Equation 3.1) did not differ significantly between four-trial conditioning (CR-history 0111) and two-trial conditioning (CR-history 01), nor between two-trial (CR-history 01) and single-trial conditioning (CR-history 0(1)). Hence, for honeybees that responded after the first conditioning trial a single CS-US pairing was sufficient to induce a stable and odor-specific 24h memory.

We also analyzed graded measures for 24h memory retention and discrimination based on proboscis extension durations. These measures overall confirmed our previous results (Figure 3.4 Aiii, iv and Figure 3.4 Biv). However, we found a significantly shorter proboscis extension duration to the CS after single-trial conditioning than after two-trial conditioning (Figure 3.4 Biii).

### 3.3.5 Honeybees responding early show higher 24h memory retention than those responding later

We asked whether the time-point during conditioning at which individual honeybees acquired the CR would have any effect on 24h memory retention and discriminatory power. In particular, we wanted to know whether associative learning during classical conditioning shares the dynamics of an Aha! effect-like learning process, referring to the hypothesis that animals may comprehend the causality between the CS and the US at a certain time-point, and as a consequence reliably respond to the CS in subsequent trials. Importantly, in the strict sense, this implies equal CR stabilities across animals with different first CR latencies. This hypothesis was formalized by a two-state hidden Markov model in a previous study (Pamir et al., 2011). Here, we explicitly compared memory retention and discrimination in animals that did not receive further training after their first CR during conditioning, assuming that the first CR indicates that an individual has switched from the naïve to the learned state (Experiment 1, Material and Methods). Both in binary and graded measures we found that animals responding early (CR-history 01) showed higher memory retention than those responding later (CR-history 001 and 0001) (Figure 3.4 Ci, iii). At the first glance, this finding argues against the Aha! effect hypothesis. However, we found no significant differences between memory's discriminatory power in animals responding early and those responding later (Figure 3.4 Cii, iv). The discriminatory power captures the pure associative effect of the conditioning procedure, devoid of non-associative effects such as sensitization (Tully, 1984). Together, these results indicate that early responders have a higher general response probability, while the truly associative component of learning is compatible with the dynamics of an Aha! effect-like learning process.

### 3.3.6 The effect of different training conditions on individual learning dynamics

We further asked how individual learning dynamics were affected by altered training conditions such as in trace conditioning, conditioning with massed training trials and differential conditioning.

In trace conditioning animals experienced a 4.5-second long stimulus-free gap between the CS and the US (dataset 13), whereas animals in the control group (delay conditioning) experienced a 1-second overlap between the CS and the US (dataset 14,). At the population level, this resulted in a lowered maximum of the regression curve $p_{max}$ during trace conditioning (Figure 3.2 G, H). We asked whether in this more difficult learning task the decrease was observed because individuals responded with lower probability to the CS, because fewer animals acquired a response at all, or because of a combination of both factors. We found that both a decreased CR stability and an increase in the number of non-responders were responsible for the decrease of $p_{max}$ (Table 3.2, dataset 13, 14). The same trend was

observable in 6 subgroups of dataset 15, in which the delay between the CS and US onset was systematically varied between 1 and 15 seconds (Table 3.2). On average animals started to respond earlier during delay conditioning compared to trace conditioning (Table 3.2, dataset 13, 14).

Comparing massed and spaced training conditions, we observed that massed trials at an inter-trial interval of 30 seconds (dataset 18, Figure 3.2 K) resulted in a decrease of $p_{max}$ as compared to the control group with spaced trials at an inter-trial interval of 15 minutes (dataset 19, Figure 3.2 L). We again found that this decrease resulted from a decrease of CR stability in individuals as well as from a higher percentage of non-responding animals. Under massed conditioning trials animals responding early ($t_{firstCR}$=2, 3) initially had high CR probability in subsequent trials but later on their CR probabilities were lowered, which may indicate a stimulus satiation effect.

Analyzing data from differential conditioning we found that the alternating presentation of unrewarded stimuli (CS-) during conditioning had no effect on the rapid and stable acquisition of a CR to the CS+ (dataset 20 Table 3.2, Figure 3.3).

### 3.3.7 Sucrose responsiveness correlates with learning performance

Our analysis showed that equally treated honeybees varied substantially in their learning performances both during simple absolute conditioning and more difficult conditioning tasks such as trace and differential conditioning. What was the reason for these learning differences across individuals? Several studies demonstrated that the responsiveness to sucrose correlates with learning performance in individual harnessed honeybees (Scheiner et al., 1999, 2001a,b, 2003, 2004, 2005; Behrends and Scheiner, 2012). We here reanalyzed data on olfactory and tactile conditioning in which individual responsiveness to sucrose of bees was determined prior to the conditioning phase (dataset 16, 17, see Materials and Methods). For both datasets we sorted animals into subgroups of equal size defined by 4 ranges of individual gustatory response scores (GRS). Subgroups with higher GRS reached higher asymptotes of average CR probability than subgroups with lower sucrose responsiveness (Figure 3.5 A, C, compare with Fig. 1, 2 in Scheiner et al. (2001a) in which learning performance was measured by acquisition scores). At the level of individuals we find that animals with higher sucrose responsiveness respond more persistently to the CS than animals with lower sucrose responsiveness (dataset 16, 17, Table 3.2). We also found that the percentage of non-responding animals was inversely correlated to the GRS (dataset 16, 17, Table 3.2, Figure 3.5 B, D). Both for olfactory and tactile conditioning we found that all animals started to respond within the first few conditioning trials (Figure 3.5 B, D), consistent with our previous results.

### 3.3.8 A heterogeneous Rescorla-Wagner model captures the learning dynamics of honeybees

The Rescorla-Wagner model (Rescorla and Wagner, 1972) provides a simple and yet influential theoretical account of associative learning during classical conditioning. Applying this theory in a straightforward way to any of our datasets from absolute conditioning allows us to estimate two parameters: the learning rate $\alpha$ and the US effectiveness $\lambda$ (Materials and Methods, Equation 3.4). However, given the substantial degree of heterogeneity in learning performance, these two group-average parameters provided an invalid description of associative learning within the population. Formally this can be shown by cross-validating the classical form of the Rescorla-Wagner model against extended versions that are able to account for heterogeneity in learning performance (Table 3.3).

In order to make the Rescorla-Wagner theory applicable to our type of data we implemented a simple extension of the model. Following the hypothesis by Scheiner et al. (2005) and our analysis shown in Figure 3.5 we assumed that different animals varied in their evaluation of the same physical sucrose reward, and as a consequence differed in their learning performances. We expressed the heterogeneity in US effectiveness $\lambda$ by a probability distribution $P(\lambda)$ (Materials and Methods). In our model, the parameter $\lambda$ captures several factors in an unspecified way: physical US properties such as concentration, duration, and temporal relationship to the CS (Sutton and Barto 1990); but also the subjective evaluation of these physical properties by an individual. Since we observed that animals always started to respond early during conditioning we further assumed that animals in a given dataset did not differ drastically in their speed of learning, which was expressed by a common learning rate $\alpha$. Fitting the extended Rescorla-Wagner model to data that was recorded under standard training conditions we typically found a bimodal shape of the distribution $P(\lambda)$ (Figure 3.6 B). The first peak at $\lambda = 0$ represented non-responding animals while the second peak at $\lambda = 1$ represented animals that started to respond in the second trial and kept responding in all subsequent trials.

Under changed training conditions we found that probability distributions were tilted towards the first peak (see Figure 3.6 C for trace vs. delay conditioning and Figure 3.6 D for massed vs. spaced conditioning). For a given animal the US effectiveness $\lambda$ decreased when the time between the CS and US onset was increased (trace conditioning), but also when the US was presented at shorter inter-trial intervals (massed conditioning). The estimated probability distributions $P(\lambda)$ replicated our previous data analyses showing a decrease of CR stability on the one hand and an increase of the percentage of non-responders on the other hand.

Originally, the Rescorla-Wagner model was proposed as a model that could explain various behavioral phenomena observed during the conditioning phase (Rescorla and Wagner, 1972). In the case of our data, where the CR in the test phase resembles a steady continuation of the behavior expressed during training, the extended Rescorla-Wagner model can also be employed

to explain memory retention in individuals as observed in Experiment 1 and 2. Given a high learning rate of $\alpha = 0.94$ (as in dataset 21 of Experiment 1), animals with $\lambda$ close to unity reach high values of associative strength after only a single conditioning trial (Figure 3.6 A). Hence the observation of an individual extending its proboscis to the second rewarded or unrewarded CS presentation already indicates a high and near asymptotic associative strength in that animal, irrespective of further training. This matches well with our observations in Experiment 1 and 2, where we observed that both 01 and 0111 subgroups and 0(1) and 01 subgroups showed high retention probabilities.

Finally, the extended model captured the general trend that animals responding later expressed lower CR stabilities than those responding earlier, which was observed both when comparing different data sets from absolute conditioning as well as in Experiment 1 (Figure 3.4 Ci). The observation of a first CR on the third or fourth trial is more likely to be emitted by an animal with a smaller value of $\lambda$, as compared to the observation of a first CR on the second trial which is more likely to be emitted by an animal with a value of $\lambda$ close to unity.

## 3.4 Discussion

Group-average learning curves confound three parameters of individual learning in the honeybee: the ability to learn a task (as indirectly observed in the percentage of non-responders), the latency in trial time until the first response is initiated, and the stability of subsequent responses. Learning curves implied that honeybees required at least three conditioning trials to reach asymptotic levels of responding, however these population dynamics were not supported at the individual level. The majority of animals that showed any response at all already extended their probosces in the second trial, i.e. after the experience of a single CS-US pairing. Once having responded for the first time, honeybees continued to respond with a high probability during training as well as in the retention test, irrespective of the number of experienced conditioning trials. In summary, three conclusions can be drawn: (1) The gradual rise of the group-average learning during the first few trials reflects the recruitment of individuals with a stable CR. It does not however provide evidence for a gradual performance increase in individuals. (2) The asymptote of the group-average learning curve mostly reflects a saturation in the recruitment of responding animals, but does not represent a performance maximum in individuals. (3) Memory retrieval in individuals resembles a constant continuation of the behavior expressed in the conditioning phase. Consequently, group-average memory retention mostly reflects the number of individuals that acquired a stable CR during training. It does not however provide a measure of memory strength or stability within the population.

**A distinction between single-trial and three-trial induced 24h memories is not supported by individual behavior** The effect of the number of conditioning trials on memory formation in the honeybee has been studied in behavior and biochemistry (for reviews see Menzel, 1999; Müller, 2012). The current model predicts lower 24h group-average memory retention after single-trial than after three-trial conditioning (Menzel, 1990 Fig. 9.8; Müller, 2012 Fig. 1; Menzel, 2012 Fig. 2). However, in individuals we could not find evidence that 24h retention probability or discriminatory power was enhanced after more training trials (Experiment 1 and 2). We suggest that the conflict between our findings at the individual level and the prevailing population-level model can be explained by a recruitment effect: After one conditioning trial typically at most half of the animals in a given conditioning group have acquired a stable CR, which leads to only moderate retention scores when the whole group is tested. On the other hand, three-trial conditioning typically recruits most of the animals that are able to learn the task, which then results in high group-average memory retention. The recruitment effect could also provide an explanation for a controversial finding by Sandoz et al. (1995), showing that honeybees can form a life-long memory even after a single conditioning trial. Group-average memory retention after single-trial conditioning drastically depends on the amount of recruitment after the first conditioning trial. It seems reasonable to assume that recruitment was high under the experimental conditions in this study (see Fig. 1 in Sandoz et

al., 1995), and that consequently group-average retention did not differ after single-trial and three-trial conditioning.

Different biochemical signatures between single-trial and multiple-trial induced memories have been repeatedly demonstrated in the honeybee: Long term memory formation requires PKA activation during conditioning (Fiala et al., 1999). In the antennal lobe PKA activity transiently increases with an increasing number of sucrose stimuli (Hildebrandt and Müller, 1995), and it is prolonged after multiple-trial, but not after a single-trial conditioning (Müller, 2000). Correspondingly, long-term memory can be induced artificially with single-trial conditioning when PKA activity is artificially prolonged during the CS-US pairing by photorelease of cAMP (Müller, 2000). In these studies, memory retention and PKA activity was measured across bees and thus represent the group average. Our findings are compatible with the current biochemical model of long term memory formation, if one assumes that the effectiveness of the underlying signal cascades differs across individuals. A bee which forms a long term memory after a single conditioning trial might have a more effective cAMP-PKA signaling cascade than a bee which requires multiple conditioning trials. Thus, our data suggests that an activation of PKA to a level that is sufficient for long term memory formation does not require multiple learning trials per se. This is also supported by a study showing that honeybees with a higher gustatory responsiveness had a higher baseline PKA activity than those with a lower gustatory responsiveness (Scheiner et al., 2003).

It would be highly desirable to test differences in PKA levels or other molecular signatures of memory consolidation under knowledge of individual behavior. The effect of training intensity could be elucidated by comparing animals with a CR history of 0(1) and 0111. In this context it should be noted that several recent studies in the honeybee enhanced their analysis of neuronal activity by taking into account individual behavior (Roussel et al., 2010; Rath et al., 2011; D'Albis et al., 2011).

**The extended Rescorla-Wagner model is compatible with individual behavior** Learning dynamics in individual honeybees, i.e. the temporal change in a hidden learning variable or state during training, can be estimated by employing simple descriptive models (see also Pamir et al., 2011). Formally, the eligibility of different model hypotheses, such as for example different Rescorla-Wagner-type models, can be determined by a cross-validation algorithm. Taking into account both formal (Table 3.3) and biological evidence (Figure 3.5), we proposed a minimal extension of the classical Rescorla-Wagner model, which was fully compatible with our behavioral observations. Following the hypothesis of Scheiner et al. (2005) we implemented a model in which different animals varied in learning performance due to their difference in perceived US effectiveness $\lambda$. A model-based data analysis allowed us to estimate the heterogeneity $P(\lambda)$ in a given conditioning group, and to visualize modulations in heterogeneity under altered training conditions (Figure 3.6).

In a previous study we hypothesized that associative learning in the honeybee may evolve

in an Aha! effect like manner, which we formally expressed by a two-state hidden Markov model (Pamir et al., 2011). Correspondingly, Smith et al. (2012) suggested Aha! effect-like learning dynamics for the case of appetitive differential conditioning (Fernandez et al., 2009) and for conditional withholding of the proboscis extension during discriminative punishment (Smith et al., 1991). Our data support this hypothesis, as the purely associative memory component - measured by the discrimination index - was equal in animals with late and early first CRs (Figure 3.4 Cii, iv).

**Implications for future research** Several factors, such as satiation level (Page et al., 1998; Ben-Shahar and Robinson, 2001; Friedrich et al., 2004), behavioral role (Scheiner et al., 1999, 2001b, 2003) or age (Behrends and Scheiner, 2012) have been shown to affect sucrose responsiveness in the honeybee, which in turn can affect learning performance. A sample of wild-type honeybees caught at the entrance to the hive will naturally consist of individuals that vary in several if not all of these factors. While satiation levels can be calibrated by standardized feeding routines before the start of an experiment (Friedrich et al., 2004), the sample will still contain an unknown composition of different types of foragers at different ages, which may cause unexplainable variability in learning performance when different experimental groups are compared. Experimentally one can control for the actual learning abilities of individuals in a given sample by testing honeybees for their responsiveness to sucrose prior to conditioning, or by selecting animals based on their performance in a preconditioning phase (Chandra et al., 2010). Complementary to this experimental approach, we suggest that analyzing data at the single-animal level may help to resolve the problem of inter- and intra-group variability, as often encountered in classical conditioning of the proboscis extension response (Frost et al., 2012; Matsumoto et al., 2012). Our analysis showed that an early and stable CR was the most salient and invariable behavioral feature of individual learning during standard training conditions. This finding may be exploited by explicitly studying the effect of altered training parameters or in vivo pharmacological (Schwärzel and Müller, 2006; Felsenberg et al., 2011) or epigenetic interventions (Lockett et al., 2010; Biergans et al., 2012) on this behavioral performance benchmark.

Computational models of plasticity in the insect brain have not been constrained by individual learning dynamics to date. In *Drosophila*, a long-held notion exists that the expression of behavior in individuals follows the group-average (Quinn et al., 1974), and only recently has this issue been touched on again (Chabaud et al., 2010). Consequently theoretical studies tend to rely on group-average performance, as for example observed in a final test phase after aversive classical conditioning in the T-maze (Young et al., 2011; Wessnitzer et al., 2012). Our study described several characteristics of associative learning in individual honeybees, and yet more data from different classical conditioning protocols may be shared by other laboratories. Integrating these behavioral constraints into current models of plasticity in the insect brain is the focus of ongoing research.

## 3.5 Tables and table captions

| Dataset | $N$ | $m$ | $T(h)$ | $ITI(min)$ | CS | Experimenter | Reference |
|---|---|---|---|---|---|---|---|
| 1 | 64 | 3 | 24,48 | 2 | clove oil | VA, JF, DE | Pamir et al., 2011 |
| 2 | 58 | 3 | 24 | 2 | clove oil | LM, JF, DE | Pamir et al., 2011 |
| 3 | 87 | 3 | 24 | 10 | clove oil | KBG, DE | Pamir et al., 2011 |
| 4 | 517 | 3 | 24 | 10 | clove oil | NS, DE | Stollhoff et al., 2005; Pamir et al., 2011 |
| 5 | 98 | 3 | 25 | 10 | clove oil | NS, DE | Stollhoff et al., 2005; Pamir et al., 2011 |
| 6 | 113 | 3 | 26 | 10 | clove oil | NS, DE | Stollhoff et al., 2005; Pamir et al., 2011 |
| 7 | 92 | 3 | 28 | 10 | clove oil | NS, DE | Stollhoff et al., 2005; Pamir et al., 2011 |
| 8 | 85 | 3 | 48 | 10 | clove oil | NS, DE | Stollhoff et al., 2005; Pamir et al., 2011 |
| 9 | 94 | 3 | 72 | 10 | clove oil | NS, DE | Stollhoff et al., 2005; Pamir et al., 2011 |
| 10 | 122 | 4 | 1,24 | 30 | isoamyl acetate | NKC | Pamir et al., 2011 |
| 11 | 37 | 5 | 1,24 | 30 | 6-pentadecene | NKC | Pamir et al., 20111 |
| 12 | 48 | 5 | 1,24 | 30 | 7-pentadecene | NKC | Pamir et al., 2011 |
| 13 | 95 | 6 | 0.25 | 10 | 1-nonanol | PS | Szyszka et al., 2011 |
| 14 | 75 | 6 | 0.25 | 10 | 1-nonanol | PS | Szyszka et al., 2011 |
| 15 | 281 | 6 | 0.25 | 10 | 1-octanol or 2-heptanone | PS | Pamir et al., 20111 |
| 16 | 100 | 11 | 24 | 5 | citral | RS | Scheiner et al., 2001a |
| 17 | 100 | 11 | 24 | 5 | tactile conditioning | RS | Scheiner et al., 2001a |
| 18 | 63 | 12 | none | 0.5 | hexanol | RM | Menzel et al., 2001; Pamir et al., 2011 |
| 19 | 64 | 12 | none | 15 | hexanol | RM | Menzel et al., 2001; Pamir et al., 2011 |
| 20 | 120 | 6 | 1,24 | 14 | 1-hexanal,1-octanol | NKC | Pamir et al., 2011 |
| 21 | 118 | 4 | 24 | 10 | 1-hexanol or 1-nonanol | PS | Unpublished data |
| 22 | 335 | 2-4 | 24 | 10 | 1-hexanol or 1-nonanol | PS | Unpublished data |
| 23 | 121 | 2 | 24 | 10 | 1-hexanol or 1-nonanol | PS | Unpublished data |
| 24 | 118 | 1 | 24 | 10 | 1-hexanol or 1-nonanol | PS | Unpublished data |

**Table 3.1 Overview over analyzed data from classical conditioning of the proboscis extension response.** Abreviations: number of animals ($N$), number of conditioning trials in the acquisition session ($m$), time of the retention test in hours after the end of the conditioning session ($T$), inter-trial-interval during conditioning trials in minutes $ITI$, conditioned stimulus (CS), Experimenters: Victoria Antemann (VA), Johannes Felsenberg (JF), Dorothea Eisenhardt (DE), Katrin Barbara Gehring (KBG), Laura Morgenstern (LM), Nicola Stollhoff (NC), Neloy Kumar Chakroborty (NKC), Ricarda Scheiner (RS), Randolf Menzel (RM), Paul Szyszka (PS); See Materials and Methods for details on individual data sets.

| Dataset | $N$ | $CR_{stability}$ (%) | $CR_{stability}(t_{firstCR}=2)$ (%) | $mean(t_{firstCR})$ (%) | $N_{non-responders}/N$ (%) |
|---|---|---|---|---|---|
| 1 | 64 | 82.5 (40) | 80.8 (26) | 2.8 | 17.2 |
| 2 | 58 | 95.6 (45) | 94.3 (35) | 2.5 | 8.6 |
| 3 | 87 | 91.7 (60) | 87.9 (33) | 2.8 | 8 |
| 4 | 517 | 87.8 (389) | 91.8 (261) | 2.6 | 9.1 |
| 5 | 98 | 90.3 (77) | 91.7 (54) | 2.6 | 7.1 |
| 6 | 113 | 92.6 (81) | 96.0 (50) | 2.8 | 7.1 |
| 7 | 92 | 87.9 (66) | 90.0 (40) | 2.7 | 9.8 |
| 8 | 85 | 90.1 (71) | 93.9 (49) | 2.5 | 4.7 |
| 9 | 94 | 94.2 (69) | 97.5 (40) | 2.7 | 8.5 |
| 10 | 122 | 70.9 (86) | 77.8 (45) | 2.9 | 20.5 |
| 11 | 37 | 81.9 (23) | 91.2 (17) | 2.3 | 37.8 |
| 12 | 48 | 81.5 (37) | 86.5 (26) | 2.7 | 16.7 |
| 13 | 95 | 52.2 (57) | 54.8 (23) | 3.5 | 35.8 |
| 14 | 75 | 83.1 (62) | 82.9 (35) | 2.8 | 14.7 |
| 15, 1s | 35 | 76.5 (26) | 80.0 (6) | 3.7 | 20 |
| 15, 2s | 34 | 65.0 (17) | 70.0 (4) | 3.8 | 44.1 |
| 15, 3s | 34 | 62.9 (19) | 80.0 (4) | 3.8 | 44.1 |
| 15,6s | 42 | 71.8 (10) | 60.0 (1) | 3.6 | 73.8 |
| 15, 10s | 34 | 37.1 (7) | 80.0 (2) | 4 | 76.5 |
| 15, 16s | 31 | 34.5 (10) | 31.4 (7) | 3 | 64.5 |
| 16, GRS 10 | 33 | 91.8 (31) | 95.6 (25) | 2.2 | 6.1 |
| 16, GRS 8-9 | 25 | 69.1 (17) | 80.8 (11) | 2.5 | 32 |
| 16, GRS 5-7 | 22 | 73.8 (16) | 100.0 (5) | 3 | 27.3 |
| 16, GRS 2-4 | 20 | 37.0 (3) | 37.0 (3) | 2 | 85 |
| 17, GRS 10 | 42 | 86.4 (40) | 95.6 (20) | 3 | 4.8 |
| 17, GRS 8-9 | 29 | 77.6 (21) | 82.2 (5) | 3.5 | 27.6 |
| 17, GRS 5-7 | 16 | 73.4 (8) | 94.4 (2) | 4.2 | 50 |
| 17, GRS 2-4 | 13 | - (0) | - (0) | - | 100 |
| 18 | 63 | 52.6 (42) | 58.3 (24) | 3 | 33.3 |
| 19 | 64 | 77.9 (51) | 91.3 (15) | 4 | 18.8 |
| 20, CS+ | 98 | 93.7 (87) | 97.0 (61) | 2.8 | 6.1 |
| 21 | 118 | 88.4 (85) | 91.7 (64) | 2.6 | 20.3 |

**Table 3.2 Summary of estimated parameters describing the dynamics of associative learning in individuals.** $N$: Number of animals in each dataset. The parameter $CR_{stability}$ equals the mean probability to respond in subsequent trials, given that animals have started to respond in any of the trials. Numbers in brackets denote numbers of responding animals. The parameter $CR_{stability}(t_{firstCR}=2)$ denotes the mean probability to respond in subsequent trials, given that animals have started to respond in the second conditioning trial. Numbers in brackets indicate the numbers of animals starting to respond in the second trial. $mean(t_{firstCR})$: Mean time-point in trial time at which animals display their first CR, given that animals respond in any of the trials. $N_{non-responders}/N$: Proportion of animals that do not respond in any of the trials. For dataset 15 (trace conditioning) parameters were computed independently for different delays between CS and US onset as indicated. For datasets 16 and 17 parameters were computed for 4 subgroups defined by gustatory responsiveness scores (GRS). For dataset 20 parameters were computed for CS+ trials only.

| **Dataset** | **Mean negative log-likelihood on test data** | | | **$\alpha$** | **$\lambda$** |
|---|---|---|---|---|---|
| | $RW$ | $RW^{P(\lambda)}$ | $RW^{P(\alpha)}$ | $RW^{P(\lambda)}$ | $RW^{P(\alpha)}$ |
| 10 | 82.6 (0) | 73.3 (2) | 73.5 (1) | 0.77 | 0.86 |
| 11 | 32.2 (0) | 18.3 (2) | 19.0 (1) | 0.98 | 0.98 |
| 12 | 40.6 (0) | 30.9 (2) | 31.3 (1) | 0.92 | 0.91 |
| 13 | 86.6 (0) | 72.3 (2) | 74.4 (1) | 0.74 | 0.74 |
| 14 | 69.4 (0) | 53.2 (1) | 52.6 (2) | 0.72 | 0.93 |
| 15, 1s | 31.4 (0) | 24.5 (1) | 23.0 (2) | 0.48 | 1.0 |
| 15, 2s | 29.0 (0) | 20.8 (1) | 20.0 (2) | 0.49 | 1.0 |
| 15, 3s | 29.6 (0) | 21.8 (1) | 21.3 (2) | 0.48 | 1.0 |
| 15,6s | 25.3 (0) | 18.0 (1) | 17.8 (2) | 0.36 | 1.0 |
| 15, 10s | 16.7 (0) | 12.8 (2) | 14.1 (1) | 0.86 | 0.86 |
| 15, 16s | 18.3 (1) | 14.3 (2) | 18.3 (0) | 1.0 | 0.64 |
| 16, GRS 10 | 40.0 (0) | 25.8 (2) | 30.6 (1) | 0.94 | 0.94 |
| 16, GRS 8-9 | 47.7 (0) | 26.4 (2) | 31.0 (1) | 0.99 | 0.81 |
| 16, GRS 5-7 | 39.5 (0) | 22.1 (2) | 23.3 (1) | 0.63 | 1.0 |
| 16, GRS 2-4 | 16.2 (0) | 9.8 (2) | 12.9 (1) | 1.0 | 0.61 |
| 17, GRS 10 | 59.0 (0) | 41.2 (2) | 43.2 (1) | 0.72 | 0.95 |
| 17, GRS 8-9 | 53.8 (0) | 30.5 (2) | 31.8 (1) | 0.52 | 0.97 |
| 17, GRS 5-7 | 29.5 (0) | 13.7 (1) | 13.2 (2) | 0.52 | 1.0 |
| 17, GRS 2-4 | 0.0 (2) | 1.8 (0) | 1.2 (1) | 1.0 | 1.0 |
| 18 | 108.8 (0) | 70.6 (2) | 80.8 (1) | 0.9 | 0.77 |
| 19 | 119.2 (0) | 71.4 (1) | 70.8 (2) | 0.52 | 0.98 |
| 21 | 76.5 (0) | 59.8 (1) | 59.0 (2) | 0.94 | 0.94 |
| SCORE | 3 | 34 | 29 | | |

**Table 3.3 Performance comparison of three Rescorla-Wagner-type models.** The performance of the classical Rescorla-Wagner model (RW) was compared with two extended versions that assumed a heterogeneity in US effectiveness $\lambda$ ($RW^{P(\lambda)}$), or a heterogeneity in learning rate $\alpha$ ($RW^{P(\alpha)}$). For each model and data set, the mean negative log-likelihood on the test data was calculated after 500 rounds of four-fold cross-validation. Numbers in parenthesis indicate the performance rank of each model for a given dataset, where 0 indicates worst and 2 indicates best model performance. For the two extended Rescorla-Wagner models, the two rightmost columns display the common learning rate $\alpha$ or US effectiveness $\lambda$ respectively, estimated on the whole dataset. Only datasets in which the maximum of the regression curve $p_{max}$ was obtained during the conditioning session were taken into account.

## 3.6 Figures and figure captions

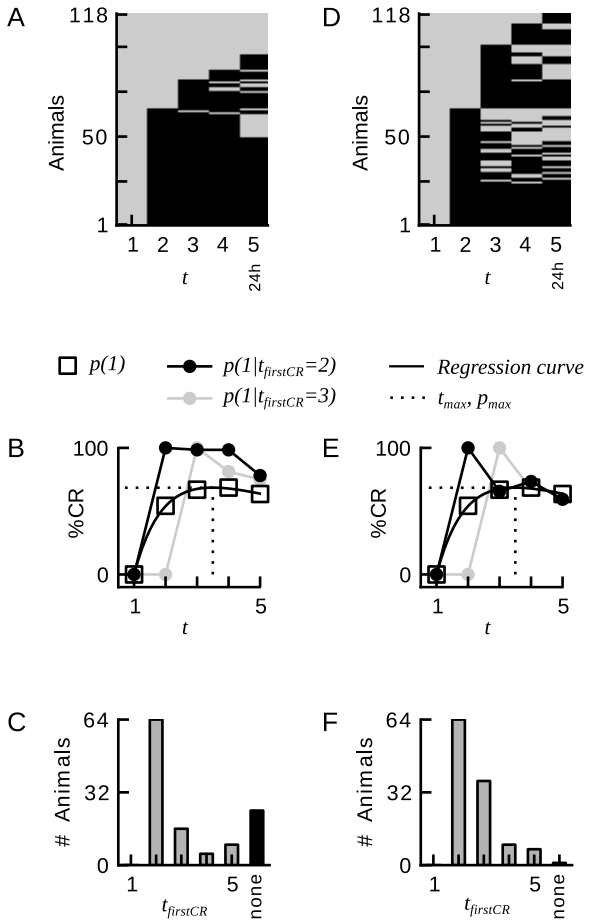


**Figure 3.1** Group-average CR probabilities do not adequately represent the CR probabilities in individual honeybees during classical conditioning of the proboscis extension response. **A** Binary conditioned response matrix from a typical dataset consisting of four conditioning trials and one memory retention test at 24h (dataset 21). A gray entry indicates no CR, a black entry indicates a CR. B Average CR probability $p(1)$ and conditional CR probabilities $p(1|t_{firstCR}{=}2)$ and $p(1|t_{firstCR}{=}3)$. Once animals have initiated their first response, they remain responding in subsequent trials with high probability. The dotted line indicates the time point in trial time $t_{max}$ at which the regression curve (Equation 3) on the average CR probabilities assumes its maximum. **C** Histogram of first responses. The largest proportion of animals starts to respond on the second trial. **D** Binary conditioned responses matrix of a hypothetical dataset, which was generated by randomly permuting the CRs of dataset 21 across animals for each trial separately. **E, F** Analog analysis to B, C. Group-average behavior represents individual behavior in the hypothetical dataset. Conditional probabilities do not reveal a serial dependency. The percentage of non-responders is drastically reduced.

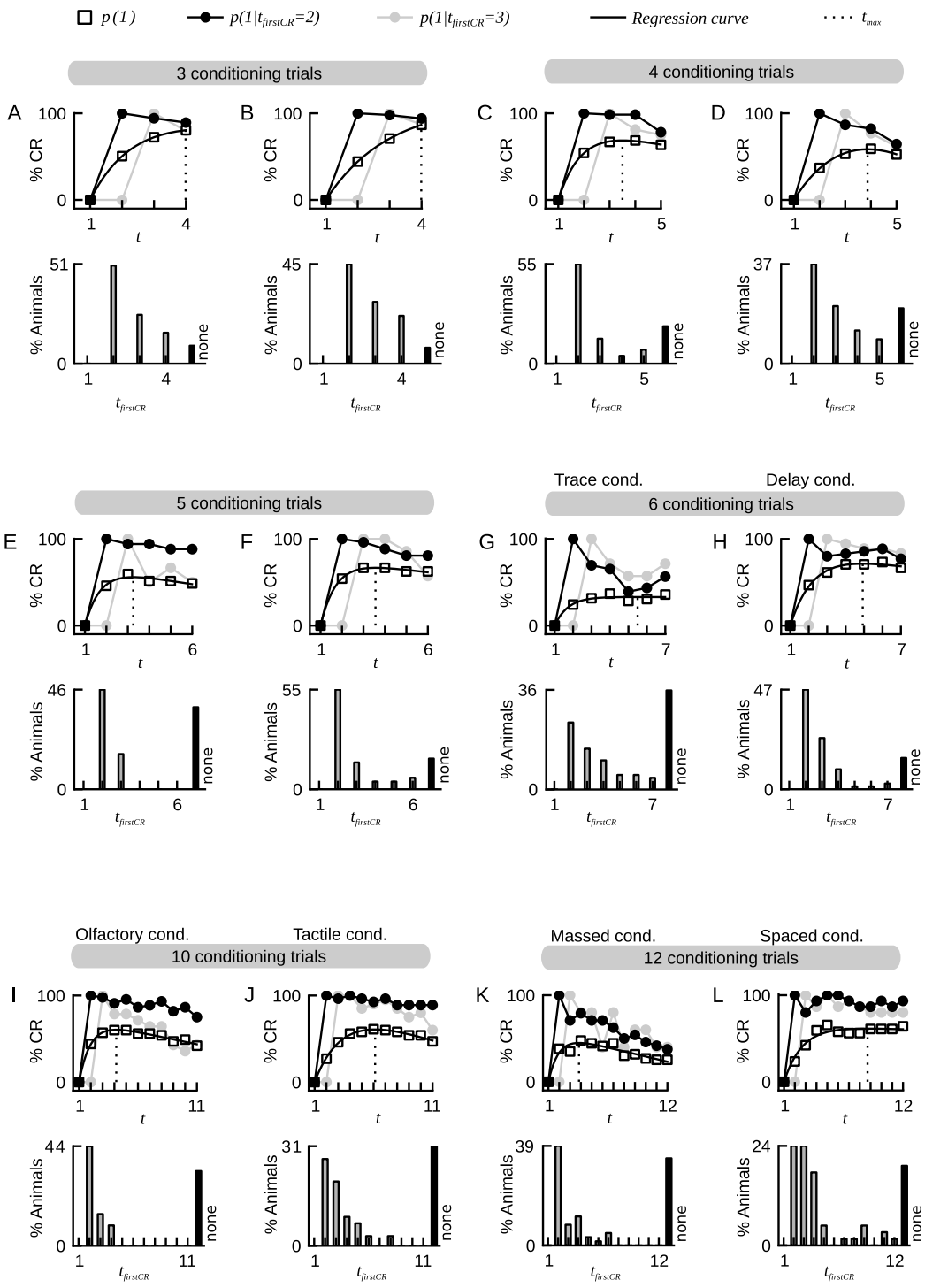


**Figure 3.2** Fast dynamics of associative learning during classical conditioning of the proboscis extension response. [*Figure caption continues on next page.*]

[*Continuation of Figure 3.2 caption*] For each data set the upper panel shows the average CR probabilities and the CR probabilities in two subgroups of animals that start to respond on the second ($t_{firstCR}$=2) or third trial ($t_{firstCR}$=3). The black line depicts a regression curve (Equation 3) on the average CR probabilities (open square symbols). The dotted line depicts the position $t_{max}$ of the maximum of the regression curve in trial time. The lower panel displays the percentage of animals that showed their first CR in a given trial. Animals that did not show a CR in any of the trials are represented by the black bar (none). Across all data sets, the largest proportion of animals starts to respond after only a single conditioning trial. Once animals have responded for the first time they have a high probability to continue responding in subsequent trials. The percentage of non-responding animals varies across datasets. Bees which responded to the first CS before the CS-US pairing were excluded from the analysis. **A** Dataset 4. **B** Dataset 6. **C** Dataset 21. **D** Dataset 10. **E** Dataset 11. **F** Dataset 12. **G** Dataset 13: The CR stability is decreased under trace conditioning (4.5 seconds gap between CS offset and US onset). **H** Data set 14: Control group for dataset 13 (CS and US overlap by 0.5 seconds).**I, J** Datasets 16 and 17: The dynamics of olfactory associative learning resemble the dynamics of tactile associative learning. **K** Dataset 18: The stability of the CR is decreased under massed training conditions (inter-trial-interval equals 30 seconds). **L** Dataset 19: The stability of the CR is high under spaced training conditions (inter-trial-interval equals 15 minutes).

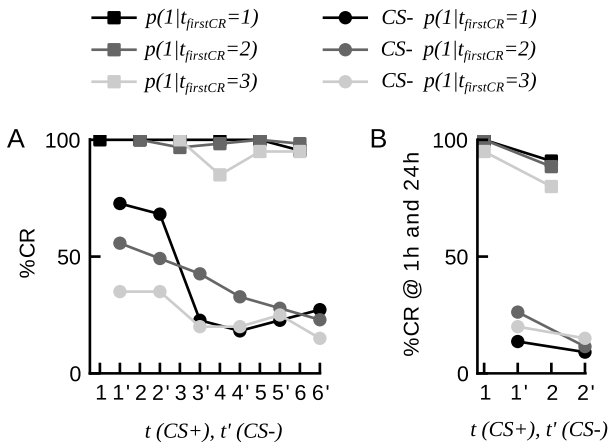


**Figure 3.3** Dynamics of discriminative learning during differential conditioning (data set 20). **A** CR probabilities to the CS+ and the CS- during the conditioning phase in three subgroups of animals, defined by their first response trial to the CS+ ($t_{firstCR}$=1, 2, 3). Once animals have started to respond to the CS+, they have a high probability to continue responding to the CS+ in consecutive trials (curves with square markers). Animals responding to the CS+ ($t_{firstCR}$=1, 2, 3) show high CR probabilities to the CS- in the first conditioning trials, and low CR probabilities to the CS- at the end of the conditioning phase. (curves with round markers). CS- trials are indicted by an apostrophe. **B** CR probabilities to the CS+ and the CS- of the three subgroups at 1h and 24h.

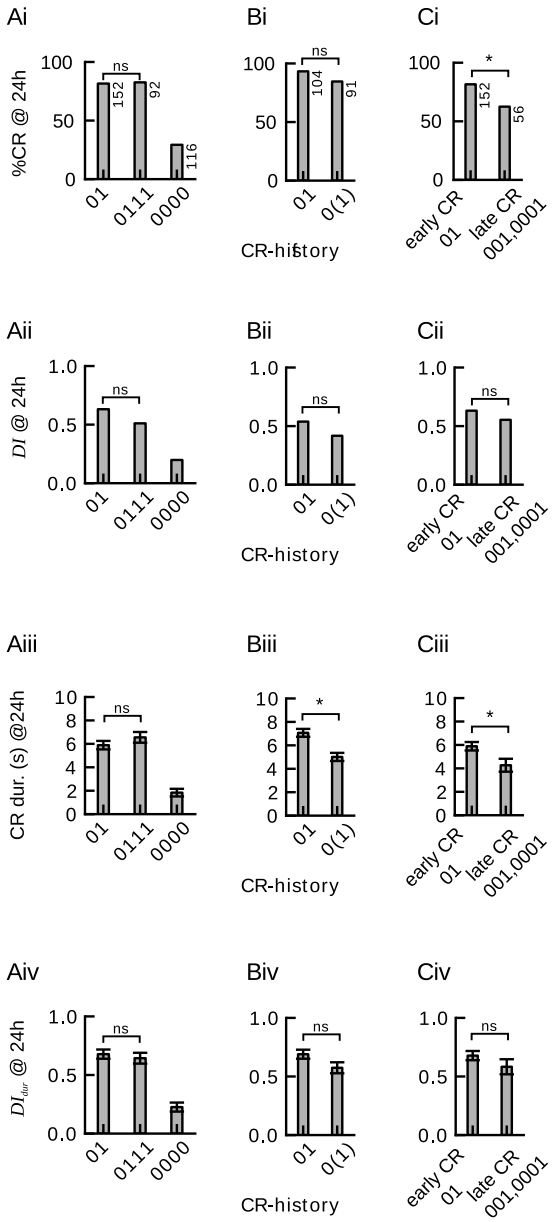


**Figure 3.4** Effect of single-trial and multiple-trial conditioning on 24h memory retention and discriminatory power under examination of individual CR histories during conditioning. [*Figure caption continues on next page.*]

[*Continuation of Figure 3.4 caption*] **Ai** CR probability to the trained odor in subgroups 01, 0111 and 0000 of Experiment 1. Memory retention after four-trial and two-trial conditioning is not significantly different (0111 vs 01 subgroup, $\chi^2 = 0.000960$ with 1 degrees of freedom, $P = 0.975$). Animals that did never respond during four-trial conditioning showed poor memory retention. **Aii** Discrimination index ($DI$) in subgroups 01, 0111 and 0000. Discriminatory power of the memory after four-trial conditioning and two-trial conditioning is not significantly different (0111 vs 01 subgroup, Mann-Whitney Rank Sum Test, $T = 10379.000$, $P = 0.095$). Animals that did not respond during four-trial conditioning show poor memory discrimination. **Aiii** Duration of the proboscis extension to the trained odor in subgroups 01, 0111 and 0000. The CR duration is not significantly different after four-trial and two-trial conditioning (0111 vs 01 subgroup, Mann-Whitney Rank Sum Test, $T = 11901.500$, $P = 0.238$). **Aiv** Discrimination Index computed on CR duration ($DI_{dur}$) in subgroups 01, 0111 and 0000. The CR duration does not reveal significant differences in memory discrimination after four-trial and two-trial conditioning (0111 vs 01 subgroup, Mann-Whitney Rank Sum Test, $T = 10652.500$, $P = 0.248$). **Bi** Memory retention after two-trial and single-trial conditioning is not significantly different (Experiment 2, 01 vs 0(1) subgroup, $\chi^2 = 2.935$ with 1 degrees of freedom, $P = 0.087$). **Bii** The discrimination index after two-trial and single-trial conditioning is not significantly different (01 vs 0(1) subgroup, Mann-Whitney Rank Sum Test, $T = 8346.000$, $P = 0.146$). **Biii** The duration of the proboscis extension response to the trained odor is significantly longer after two-trial than after single-trial conditioning (01 vs 0(1) subgroup, Mann-Whitney Rank Sum Test, $T = 7265.500$, $P =< 0.001$). **Biv** The duration of the proboscis extension response did not reveal significant differences in discriminatory power after two-trial and single-trial conditioning (01 vs 0(1) subgroup, Mann-Whitney Rank Sum Test, $T = 8256.500$, $P = 0.093$). **Ci** Animals that started to respond early during conditioning showed significantly more memory retention than animals that started to respond later during conditioning (Experiment 1, 01 vs (001, 0001) subgroup, $\chi^2 = 7.246$ with 1 degrees of freedom, $P = 0.007$). **Cii** Early and late responders do not significantly differ in memory discrimination (Mann-Whitney Rank Sum Test, $T = 5489.000$, $P = 0.346$). **Ciii** The duration of the proboscis extension response to the trained odor is significantly longer in early than in late responders (Mann-Whitney Rank Sum Test, $T = 4925.500$, $P = 0.016$). **Civ** The duration of the proboscis extension response does not reveal significant differences in memory discrimination between early and late responders (Mann-Whitney Rank Sum Test, $T = 5428.000$, $P = 0.271$).

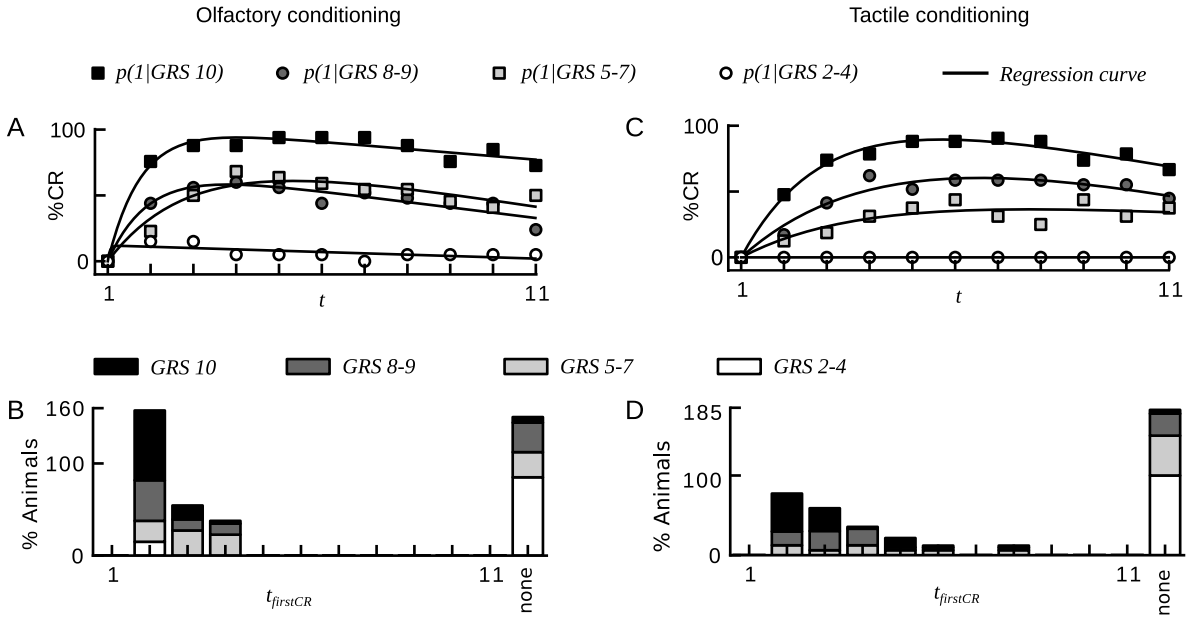


**Figure 3.5** Sucrose responsiveness correlates with learning performance in olfactory and tactile classical conditioning (compare to Figs. 1 and 2 in Scheiner et al., 2001a). **A** Average CR probabilities in four subgroups of animals from olfactory conditioning (dataset 16). Animals were divided into subgroups on the basis of individual gustatory response scores (GRS). Small numbers indicate low responsiveness to sucrose. Animals with high responsiveness to sucrose reach higher plateaus in CR probability than animals with low responsiveness to sucrose. **B** Histogram showing the percentage of animals for each subgroup that start to respond in a given trial. Animals showing at least one CR in any of the trials start to respond early during conditioning. Most non-responders have a low responsiveness to sucrose. **C** Average CR probabilities in four subgroups of animals from tactile conditioning (dataset 17). **D** Histogram showing the percentage of animals for each subgroup that start to respond in a given trial. The dynamics of tactile learning resemble the dynamics of olfactory learning.

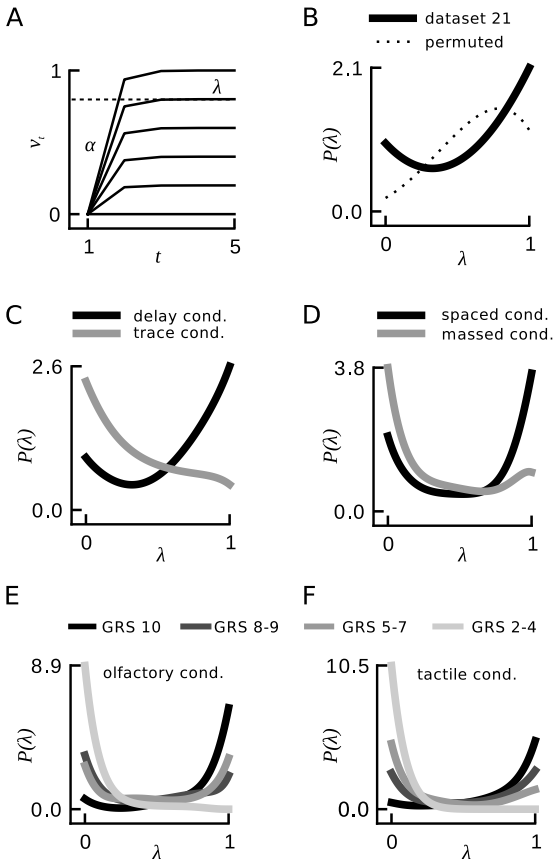


**Figure 3.6** An expanded Rescorla-Wagner model can capture the dynamics of associative learning in individual animals during classical conditioning. **A** Across conditioning trials animals gradually learn the association between the conditioned stimulus (CS) and the sucrose reward (US). The effectiveness of the US $\lambda$ differs across animals and hence different animals reach different plateaus of associative strength (AS). All animals are assumed to learn at the same speed. For the depicted example traces the learning rate equals $\alpha = 0.94$ (as in dataset 21). The observed CR probability in individuals is assumed to directly reflect the associative strength. **B** The heterogeneity in US effectiveness is estimated by a probability distribution $P(\lambda)$ (see Materials and Methods). The typical bimodal shape (B) of the probability distribution is tilted under altered training conditions (C-F). The distribution assumes the shape of a normal distribution if the CR histories at each trial are permuted across animals (dotted line for dataset 21, $\lambda_{mean} = 0.79$, compare with permuted CR matrix in Figure 1D). **C** Probability distribution for trace (dataset 13) and delay conditioning (dataset 14). **D** Probability density function under massed (dataset 18) and spaced conditioning trials (dataset 19). **E** Probability distributions in four subgroups of animals from olfactory conditioning (dataset 16). Animals have been grouped into subgroups on the basis of individual gustatory response scores (GRS). Small numbers indicate low responsiveness to sucrose. **F** Probability density function in four subgroups of animals from tactile conditioning (dataset 17).

# Chapter 4

# From neuronal computation to behavioral plasticity: A model-based investigation of associative learning in honeybees

**Abstract**

Honeybees possess a remarkable repertoire of associative learning faculties, many of which have been characterized by classical conditioning of the proboscis extension response. Behavioral studies were accompanied by numerous physiological studies that provided insights into olfactory information processing and reward-dependent plasticity in the honeybee brain. Despite the wealth of data on behavioral and neuronal plasticity in this animal model, no explicit link has yet been made between the two by a computational model or theory. We here collected several key findings on behavioral plasticity as observed over consecutive training trials in various different absolute and non-absolute classical conditioning protocols. We try to explain these learning dynamics by a set of biologically motivated computations in a simple circuit model of the honeybee brain. We present a basic model hypothesis which is compatible with behavior for most simulated conditioning protocols. This model provides insights into the effect of external stimulus properties, internal stimulus encoding schemes, and computational principles such as divisive normalization or associative learning rules on the emergence of behavior over conditioning trials in the honeybee. Our study defines a framework for the quantitative comparison of different model hypotheses on the basis of a large collection of trial-resolved behavioral data. This framework can be employed in future studies to evaluate computational models that are implemented at higher degrees of biological realism.

## 4.1 Introduction

The honeybee brain hosts a remarkable repertoire of associative learning faculties, and the link to cognition in higher animals has been repeatedly made (Giurfa, 2003; Menzel et al., 2007; Menzel, 2012). Learning under various experimental conditions has been characterized extensively by olfactory conditioning of the proboscis extension response (Bitterman et al., 1983; Giurfa and Sandoz, 2012; Matsumoto et al., 2012). In this classical conditioning paradigm, honeybees have to learn that an odorant (conditioned stimulus CS) is followed by a sucrose reward (unconditioned stimulus US). The learning performance during training is monitored by computing the group-averaged probability for a proboscis extension (conditioned response CR) on each conditioning trial. The multitude of different variants of this protocol can be roughly categorized into absolute and non-absolute conditioning protocols. In absolute conditioning, honeybees are presented to only one conditioned stimulus during training. The effect of several experimental parameters on learning can be studied in this type of protocol, such as the concentration and duration of sensory and gustatory stimuli (Pelz et al., 1997; Wright and Smith, 2004; Wright et al., 2009), the inter-stimulus interval between odor and sucrose reward (Szyszka et al., 2011) or the inter-trial-interval (Menzel et al., 2001). Non-absolute conditioning on the other hand entails learning stimulus-reward contingencies for more than one odorant during training. For example, in differential conditioning, the animal has to learn that one odorant is followed by a reward, while another is not. This task can be made more difficult by reducing the concentration of odorants, or by employing a mixture of odorants at different ratios as the rewarded and unrewarded stimulus, respectively (Fernandez et al., 2009). Even more complicated variants of this protocol introduce an ambiguity in the value of individual stimuli by presenting them both in rewarded and unrewarded trials, as in negative or positive patterning protocols (Deisig et al., 2001).

The wealth of observations on behavioral plasticity in the honeybee is accompanied by numerous neuroanatomical, biochemical and physiological studies on olfactory information processing and learning in the honeybee circuitry (Sandoz, 2011; see also Table 1 in Himmelreich and Grünewald, 2012, for a list of learning-related events in the honeybee brain). Despite the huge amount of data on trial-resolved behavioral plasticity on the one hand, and neuronal information processing on the other, no theoretical attempt has been made so far to explicitly link the two by a computational model. We here try to explain the observed changes in conditioned response probabilities over training trials by neuronal computation in a simple network model of the honeybee brain. Our study builds on recent modeling work in the insect (Huerta and Nowotny, 2009), in particular in the fruit fly (Young et al., 2011; Wessnitzer et al., 2012). In the fly, behavioral plasticity in individuals over training trials is typically not monitored, and consequently models are only constrained by group-averaged behavioral performance scores measured after training (Young et al., 2011; Wessnitzer et al., 2012). The case is different for classical conditioning in harnessed honeybees, where the experimental procedure

allows to follow and record the behavior of each individual over training trials (Felsenberg et al., 2011; Matsumoto et al., 2012). The existence of raw data at the level of individuals also provided the basis for studying inter-individual differences in learning performance in the honeybee (Scheiner et al., 2001b, 2004; Pamir et al., submitted manuscript). As we will show, knowledge on group heterogeneity is essential for the correct definition of behavioral model constraints. In Drosophila, inter-individual differences in learning performance are typically not considered by theoretical studies because individual behavior is not recorded, but also because the expression of behavior in a sample of fruit flies is assumed to be homogeneous (Quinn et al., 1974).

Our study concentrates on several key findings on behavioral plasticity in the honeybee (Table 4.1), which were observed under various different training protocols (Table 4.2). We aim at a minimal model of biologically motivated computations that can reproduce or potentially explain the behavioral dynamics of learning in these data sets. In terms of model complexity we implement a simple information processing scheme along the sensor-to-motor circuitry in the honeybee, which allows us to investigate the effects of stimulus parameters, network geometry, and computational principles such as divisive normalization or associative learning rules on the emergence of behavior (Carandini, 2012). Future model extensions towards a full-blown spiking neuronal network implementation of the current model are discussed in the discussion section. The following methods section describes the settings of the minimal model, also referred to as the basic model. The results section investigates how well this model fits to behavioral data.

## 4.2 Model methods

### 4.2.1 Behavioral model constraints

Classical conditioning of the proboscis extension response in the honeybee was introduced more than fifty years ago (Takeda, 1961; Giurfa and Sandoz, 2012). Details on this experimental procedure can be found elsewhere (Scheiner et al., 2001b; Felsenberg et al., 2011; Matsumoto et al., 2012). We collected a large set of classical conditioning protocols from the experimental literature (Table 4.2).

For some of these experimental conditions the experimental raw data at the level of individual animals was available. For these cases we computed and plotted the average CR probabilities over training trials in order to obtain a target trace for the simulations (Figure 4.6 black curves). In some of these datasets (absolute, massed, delay, differential, negative patterning A+) we excluded non-responding animals, which results in a more homogeneous sample of animals with respect to learning performance (Pamir et al., submitted). This procedure was not applicable to datasets in which the learning task was made more difficult by increasing the inter-stimulus interval (trace 1 second, 5 seconds, 10 seconds, data from Szyszka2011). In these datasets we subtracted a hypothetical basic level of non-responders, which was estimated based on the delay conditioning group.

In the datasets in which the gustatory response score of individual animals was determined prior to conditioning (absolute with GRS, data from Scheiner et al., 2001b), we did not exclude any non-responders. (In this case, a specific correlation exists between the GRS and the percentage of non-responders. See also discussion.)

For some conditioning protocols the raw data was not available (latent inhibition, differential conditioning at low concentratins and mixtures, negative and positive patterning). To obtain target traces for the simulations we transcribed the CR probabilities manually from the original studies (Fernandez et al., 2009; Deisig et al., 2001; Chandra et al., 2010).

### 4.2.2 Model geometry

The geometry of the basic model is depicted in Figure 4.2, and a summary of all model parameters is listed in Table 4.3. We used a simple three-layer feed-forward network model with excitatory connections only: 49 projections neurons (PN) in the antennal lobe connect to 5000 Kenyon cells (KC) in the Mushroom body via the two-dimensional connection matrix $c$ of size $5000 \times 49$. The probability for a connection between a PN and a KC was set to $p_{PN \to KC} = 0.204$, which ensures that each KC receives input from 10 PNs. This resembles experimental observations in the fruit fly (Turner et al., 2008). We employed a fixed and pseudo-randomized connection matrix (Fig. 4.3 B) such that each KC received input from exactly 10 projection neurons, and each PN connected to exactly 1020 KCs. Connection strengths were set to unity. Employing a fixed connection matrix removed all stochastic

components from our model, and hence no repeated simulation runs were required for a given conditioning protocol.

All Kenyon cells converge on a single extrinsic neuron via the synaptic weight vector $w$ of size 5000. Initial weights were set to zero and underwent changes during classical conditioning. Al Kenyon cells receive a global reinforcement signal, which mimics the activity of a giant octopaminergic neuron in the honeybee brain that signals reward to the Mushroom bodies as well as other regions in the honeybee brain (Hammer, 1993; Kreissl et al., 1994).

### 4.2.3 From antennal lobe input to motor output

Our mathematical formulation for the processing and transformation of olfactory information in the insect brain builds upon previous modeling work by Huerta and Nowotny (2009), Schmuker et al. (2011), Young et al. (2011), and Wessnitzer et al. (2012). For studying the effect of external stimulus properties on behavioral performance we made use of three toy olfactory stimuli of size 49, referred to as A, B and AB (Figure 4.3 A). Stimuli are constructed by a sinus function on the interval $[1, 49]$. The $i$-th entry of stimulus A is computed as

$$A_i(t) = a(t)\left[\sin\left(\frac{2\pi\ i)}{49}\right)\right]^+, \tag{4.1}$$

where $a(t)$ equals the temporal trace of stimulus A, as defined by the parameters stimulus onset, duration and intensity (see Table 4.2 and Figure 4.3 A). The parameter $overlap \in [0, 1]$ allows adjusting the similarity between stimulus A and B (Figure 4.3 A). The $i$-th entry of stimulus B is computed by shifting the phase of the sinus function:

$$B_i(t) = b(t)\left[\sin\left(\frac{2\pi\, i - overlap\, \pi)}{49}\right)\right]^+, \tag{4.2}$$

where $b(t)$ equals the temporal trace of the presentation of stimulus B. The compound stimulus AB is constructed by adding A and B.

Experimental observations in Drosophila showed that the activation of PNs saturates as a function of increasing firing in the olfactory receptor neurons in the antenna (Bhandawat et al., 2007; Kazama and Wilson, 2008). As in modeling work by Schmuker et al. (2011) we account for this behavior by computing the activation of the $i$-th glomerulus in the antennal lobe $GL_i$ by a logarithmic transfer function:

$$GL_i(t) = \ln(1 + X_i(t)), \tag{4.3}$$

where $X$ equals the sum of all stimuli that are present at time $t$. In the case of the compound stimulus AB, this corresponds to assuming a linear superposition of stimuli at the olfactory receptor level, which is in line with experimental observations in the honeybee (Deisig et al., 2006; see also Schmuker et al., 2011).

At each simulation time step $t$, the activation of the $i$-th glomerulus normalized by the total glomerular activation defines the activation of the $i$-th projection neuron:

$$PN_i(t) = \frac{GL_i(t)}{\sqrt{\sum_{j=1}^{N_{GL}} GL_j(t)^2}} \qquad j = 1, ..., N_{PN} \tag{4.4}$$

We here employed the same normalization rule as Wessnitzer et al. (2012). Activity of the i-th Kenyon cell $KC_i(t)$ is defined as:

$$KC_i(t) = \Theta\Big(\sum_{j=1}^{N_{PN}} c_{ij} PN_j(t) - \theta_{KC}\Big) \qquad i = 1, ..., N_{KC} \tag{4.5}$$

where $c$ is the 2-dimensional connection matrix of size $(N_{KC} \times N_{PN})$, $\theta_{KC}$ is the Kenyon cell activity threshold, and $\Theta(x)$ is the Heaviside function with:

$$\Theta(x) = \begin{cases} 1 & \text{if } x \geq 0; \\ 0 & \text{if } x < 0. \end{cases} \tag{4.6}$$

The chosen Kenyon cell activity threshold of $\theta_{KC} = 1.5$ ensures that on average $(2-5)\%$ of all Kenyon cells are active for any stimulus presentation which resembles experimental observations (Wilson et al., 2004; Turner et al., 2008). The activity in the single extrinsic neuron (EN) equals:

$$EN(t) = f\Big(\sum_{i=1}^{N_{KC}} w_i KC_i(t)\Big), \tag{4.7}$$

where $w_i \in [0, 1]$ is the synaptic weight of the $i$-th KC, and $f$ implements a linear threshold function for the EN activation:

$$f(x) = \begin{cases} x/168 & \text{if } x/168 \leq 0.95; \\ 0.95 & \text{if } x/168 > 0.95. \end{cases} \tag{4.8}$$

For the chosen parameters the argument of the transfer function $f$ will be below threshold for most simulated conditioning protocols. Our choice of $f$ readily allows taking the EN activation as a measure of the conditioned response (CR) probability on trial $T$:

$$p^{CR}(T) = \langle EN(t) \rangle^{observation\ window}, \tag{4.9}$$

The observation window refers to the time interval in trial $T$ in which the experimenter records the CR, typically the CS presentation not overlapping with the US. The time-step for all simulations was set to 0.1 seconds.

### 4.2.4 Reward-dependent learning in the mushroom-body output

For the basic model we assume that synaptic plasticity primarily depends on two factors: (1) the amount and timing of KC activation, which is expressed by an eligibility trace, and (2) the presence of a global reinforcement signal at the KC synapses. The eligibility trace $e_i(t) \in [0, 1]$ of the $i$-th synapse is implemented as

$$e_i(t+1) = e_i(t) - e_i(t)/\tau_e + \alpha_e \, KC_i(t) \, (1 - e_i(t)), \tag{4.10}$$

where $\alpha_e$ is the growth rate, and $\tau_e$ is the decay constant of the eligibility trace. Our formulation of an eligibility trace that holds the information about the odorant stimuli in the KCs until the arrival of the reward signal deviates from the mechanism proposed by Wessnitzer et al. (2012). In their spiking neuronal network model, the eligibility trace (or synaptic tag) was set depending on both pre- and postsynaptic activity via spike-timing-dependent plasticity. In our formulation, the eligibility of a KC to undergo changes in synaptic efficacy depends on KC activity only. Izhikevich (2007) pointed out several molecular mechanism that may implement an eligibility trace, and Yarali et al. (2012) proposed a detailed biochemical model which could explain the effect of inter-stimulus-interval on behavioral performance in Drosophila.

The octopamine concentration at the synapse $OCT(t) \in [0, 1]$, which signals reinforcement, is defined as

$$oct(t+1) = oct(t) - oct(t)/\tau_{oct} + \alpha_{oct} \, S \, R(t) \, (1 - oct(t)), \tag{4.11}$$

where $\alpha_{oct}$ is the growth rate, $\tau_{oct}$ is the decay constant of the octopamine concentration, $R(t) \in [0, 1]$ is the sucrose stimulus as defined by the simulated classical conditioning protocol, and $S$ is a linear factor that represents the subjective evaluation of the sucrose reward by a given honeybee (Scheiner et al., 2004).

On the basis of the eligibility trace and the reinforcement signal we compute the triggered short-term memory (STM) target weight $w_i^{target}(T)$ for the $i$-th synapse as

$$w_i^{STM}(T) = 0.04 \sum_{t=trial\,start}^{trial\,end} e_i(t) \, oct(t), \tag{4.12}$$

The factor 0.04 ensures that $w^{STM} \in [0, 1]$ for all simulated conditioning protocols (given the simulation time step equals 0.1 seconds). The formulation of a target weight expresses the finding that weight changes induced on a given conditioning trial require a certain period time to consolidate (Menzel, 2001). Hence, the target rate depends on the inter-trial-interval (ITI), which is modeled by a linear function

$$f^{STM}(x) = \begin{cases} x \, \mathrm{ITI}/36\,s & \text{if ITI} < 36\,s \\ x & \text{if ITI} \geq 36\,s \end{cases} \tag{4.13}$$

Finally, after each conditioning trial $T$, the $i$-th synaptic weight $w_i \in [0,1]$ is changed according to the learning rule

$$w_i(T+1) = w_i(T) + \alpha_w \Big[ f^{STM}\Big( w_i^{STM}(T) \Big) - w_i(T) \Big]^+, \tag{4.14}$$

where $\alpha_w \in [0,1]$ is the learning rate.

### 4.2.5 Parameter search and model comparison by an error function

We implemented a simple brute search parameter optimization algorithm that minimizes the squared errors $e^2$ between the observed and the simulated conditioned response probabilities. For a given classical conditioning protocol $x$ and model $h$ the sum of the squared errors over trials equals

$$e^2(x,h) = \sum_{T=1}^{N_{trials}} \Big( p^{CR,\,data}(T) - p^{CR,\,sim}(T) \Big)^2. \tag{4.15}$$

In addition, the parameter search can be performed by minimizing the total squared error $E^2$ on a collection of classical conditioning protocols given a model $h$:

$$E^2(h) = \sum_x e^2(x,h). \tag{4.16}$$

Optionally, individual error terms in Equation 4.16 may be weighed depending on data size or other heuristic factors such as data credibility. Finally, the total error of a given model hypothesis $h$ is defined as the sum of the squared errors over all classical conditioning protocols that were considered in the present study (Table 4.1). In the following we abbreviate the basic model by $h_1$.

## 4.3 Results

In this section we try to establish a quantitative link between neuronal computation in the honeybee brain circuitry and the observed dynamics of behavioral plasticity over consecutive training trials. We will first configure the basic model $h_1$ on a collection of data recorded in different absolute conditioning protocols, and than study how well the model generalizes when simulating behavior in non-absolute conditioning protocols. In this process we will introduce and discuss necessary extensions of the basic model.

### 4.3.1 Model configuration on data from absolute conditioning

**Average network activity** We first confirm that our network configurations and employed stimuli reproduce the sparse activity of KCs in the Mushroom bodies, as reported in several insect species (Perez-Orive et al., 2002; Wilson et al., 2004; Turner et al., 2008). For the first toy stimulus A, the average activity of a PN that receives non-zero glomerular input (Equation 4.4) equals 0.19. In order to meet the KC activation threshold of $\theta_{KC} = 1.5$, a KC requires around 8 non-zero PN inputs to fire. For the presentation of stimulus A, 168 KCs are above threshold, which equals 3.4% of the total KC population. As defined by our learning rule (Equation 4.14), the synaptic weights of these 168 KCs will change during absolute conditioning. Given the fixed connectivity matrix $c$, the network activation for stimulus B is almost similar: 174 KCs are above threshold, which equals 3.9% of the total KC population. For the compound stimulus AB the number of active KCs depends on the overlap between stimulus A and B. For $overlap = 0.1$, 539 KCs are active (11.9%), for $overlap = 0.5$, 329 KCs are active (6.6%), and for $overlap = 0.9$, 179 KCs are active (3.6%).

**Fitting the parameters $\alpha_w$ and $\tau_e$** Given our choice of model settings, the performance of the model $h_1$ during absolute classical conditioning can be adjusted by two free parameters: the learning rate for the change in synaptic weights $\alpha_w$, and the decay constant of the eligibility trace $\tau_e$. Figure 4.4 B shows the outcome of the parameter search for these two parameters on five conditioning protocols (absolute, delay, trace 1 second, trace 5 seconds, trace 10 seconds). Parameters were optimized by minimizing the sum of the normalized squared errors between simulated and observed CR probabilities for the five protocols (Equation 4.16).

Figure 4.4 A illustrates the effect of the eligibility trace ($\tau_e = 6.9\,s$) on the synaptic weight change. If the reward shortly follows the olfactory stimulus, the product of the eligibility trace and the octopamine trace will be large (Equation 4.12), which results in a high target weight in the learning rule (Equation 4.14). If however, in trace conditioning (Szyszka et al., 2011), the inter-stimulus interval between odor and reward increases, the target weight will be reduced, which then results in poorer learning. The actual difference between experimental and simulated CR probabilities for the five absolute conditioning protocols is shown in Figure 4.6 A,C,D and E.

**The basic model captures the rapid and stable acquisition of the conditioned response in individuals** Behavioral plasticity in individual honeybees during absolute classical conditioning can be characterized by two features (Pamir et al., submitted): First, animals start to respond very early during conditioning. Typically 94% of all animals that show at least one response in any of the trials start to respond within the first four conditioning trials. Second, once a honeybee has elicited its first CR, it continues to respond in subsequent trials with a high probability of around 86%. As shown in Figure 4.5, the basic model reproduces these behavioral features. It should be noted that the model assumes identical learning dynamics for all 63 simulated animals. Previous analysis showed that this may not be the case in the actual sample of trained animals, because animals that responded early ($t_{firstCR}$=2) tend to have higher response probabilities in subsequent trials than those responding later ($t_{firstCR}$=3, 4) (Pamir et al., submitted). As it can be seen in Figure 4.5 B, for high learning rates ($\alpha_w = 0.68$), these subtle inter-individual differences are less pronounced and hence neglectable (see Discussion).

**The effect of sucrose sensitivity and inter-trial interval on learning** Several studies showed that honeybees can differ in learning performance due to a differential evaluation of the sucrose reward (Scheiner et al., 2001b, 2004). We account for this behavior by an additional linear factor in Equation 4.11, referred to as the sucrose sensitivity $S$. For the basic model $h_1$, this parameter was set to unity. In some of the data, the gustatory response score (GRS) of individuals was determined prior to the conditioning phase (see Scheiner et al., 2001b for data origin and details on experimental procedures). On the basis of these datasets, our formalism allows determining a mapping between experimentally assessed GRS and reward evaluation within the circuitry. Figure 4.7 shows this mapping for three intervals of GRS. The factor $S$ has been treated as a free parameter in the parameter search (Equation 4.15). Having determined this mapping, the basic model hypotheses can reproduce the correlation between GRS value and learning performance as observed at the group-average level (Figure 4.6 B). It should be noted that the lowered asymptote observed in subgroups with lower gustatory response scores (Figure 4.6 B) is only partially caused by a decrease in CR probability in individuals. The main contribution comes from an increased percentage of non-responding animals (Pamir et al., submitted). The current model does not entail this level of detail (see Discussion).

Behavioral data from classical conditioning at short or long inter-trial intervals shows that changes in synaptic efficacy induced in a given conditioning trial require a critical time period to consolidate (Menzel, 2001). If the time interval between two training trials is too short, this process will be impaired. We implemented this mechanism by introducing a dependency of the target weight change on the inter-trial interval (Equation 4.13) in the learning rule (Equation 4.14). As a consequence, the model $h_1$ reproduces poor learning under massed conditioning trials (Figure 4.6 C) as described in Menzel (2001).

### 4.3.2 Model performance in non-absolute conditioning protocols

The previous section described the calibration of the basic model $h_1$ on absolute conditioning data. The model now captures a wide range of behavioral phenomena in a quantitative way. The effect of different model parameters and computational functions on behavior can be studied, and quantitative predictions for untested classical conditioning protocols can be made. However, the model has no explanatory power so far, because the formalism was designed to reproduce the behavioral phenomena. In the present section we investigate how well the calibrated model $h_1$ generalizes to non-absolute conditioning tasks.

**Differential conditioning** Figure 4.6 G shows the performance of the model $h_1$ for a differential conditioning protocol. Notably, the model readily captures the response probabilities for the rewarded stimulus A, although the learning rate $\alpha_w$ was determined on absolute conditioning data only. This again confirms that learning of rewarded stimuli in absolute and differential conditioning share the same fast dynamics (Pamir et al. 2011, (submitted)). In our model, generalization between the rewarded and the unrewarded stimulus depends on our choice of the connection matrix $c$ (MB encoding scheme), as well as on the overlap between the two toy stimuli. Treating the overlap as a free parameter allows fitting the amount of generalization observed in the behavioral traces (Figure 4.6 G). For $overlap = 0.8$, 100 KC will be activated exclusively by stimulus A, 105 KCs will be activated exclusively by stimulus B, and 68 KCs will be activated both by the presentation of stimuli A in rewarded trials, and by the presentation of stimulus B in unrewarded trials (Figure 4.3). (The latter population is responsible for the generalization between stimulus A and B. Changing the stimulus overlap results in a vertical shift of CR probabilities to stimulus B.)

The model so far captures the approximate degree of generalization, however it cannot explain the increase in discrimination between rewarded and unrewarded stimuli which is typically observed in differential conditioning at high odor concentrations. We now introduce two biological mechanisms that may be responsible for the observed increase in discrimination over training trials.

(1) Several studies found evidence for reward dependent plasticity during differential conditioning in the antennal lobe (Faber et al., 1999; Fernandez et al., 2009; Rath et al., 2011). We here implement a simple stimulus decorrelation mechanism by shifting stimulus B by the phase $\Delta overlap$ after every second conditioning trial. (Hence after the experience of a rewarded an an unrewarded trial.) Treating $\Delta overlap$ as a free parameter we find that a phase increment of $\Delta overlap = 0.05$ is compatible with the observed behavioral dynamics (Figure 4.8 A). (See discussion for more sophisticated implementations of reward-dependent pattern decorrelation mechanism in the antennal lobe.)

(2) As a second biological mechanism we consider the possibility that the Mushroom body stores information about unrewarded stimuli in the form of inhibitory synaptic weights. In

analogy to Equation 4.12, we implement this hypothesis by assuming that activity in the KCs which is *not* followed by a reward induces a target *inhibitory* weight in trial $T$:

$$w_i^{inhibit,\,STM}(T) = 0.017 \sum_{t=t_{start}}^{t_{end}} KC_i(t). \tag{4.17}$$

Negative weight changes are then implemented by the following rule for inhibitory learning:

$$w_i(T+1) = w_i(T) - \alpha_{inhibit}\, f^{STM}\Big(w_i^{inhibit\,,STM}(T)\Big), \tag{4.18}$$

where $\alpha_{inhibit} \in [0,1]$ is the inhibitory learning rate, and the maximum negative synaptic weight is set to -1.3. (Alternatively to this formulation with negative synaptic weights, one may implement a second inhibitory MB output pathway targeting a second extrinsic neuron that inhibits the first one.) In order to fit this extended model to the differential conditioning data, we first set the overlap between stimulus A and B to 0.9, which ensures a correct degree of generalization on trial $T = 2$ (Figure 4.8 B). We then treat both the learning rate $\alpha_w$ and the inhibitory learning rate $\alpha_{inhibit}$ as free parameters (Equation 4.16) and find that the parameter combination ($\alpha_w = 0.82,\ \alpha_{inhibit} = 0.27$) best reproduces the behavioral observations (Figure 4.8 B). Figure 4.3 C shows the KC stimulus encoding scheme for this simulation. In rewarded trials (stimulus A), a total of 168 KCs is activated and hence undergoes positive changes in synaptic weights. In unrewarded trials (stimulus B), 111 KCs that were already activated by stimulus A are again active and undergo relatively small negative changes in synaptic weight. In addition, 55 KCs which are exclusively active to stimulus B gain negative synaptic weights. The net effect is an increase in discrimination over trials as observed in experimental data.

In summary both decorrelation in the antennal lobe, and inhibitory learning in the Mushroom body are compatible with behavior. Importantly, each of the two model extensions has its own distinctive fingerprint under changed conditioning protocols. For example, the second extension can also explain a retarded acquisition after a pre-exposure to unrewarded stimuli, as observed in a study by Chandra et al. (2010), while the first extension will not account for this observation.

Figure 4.6 H and I show observed and simulated behavior for differential conditioning at low odor concentrations, and with mixture stimuli. The data does not provide evidence for an increase in discrimination over training trials, and consequently none of the two mechanisms implemented above was presumably operating under these more difficult learning tasks. It should be noted that unlike for all other datasets considered so far, the percentage of non-responders for the two differential conditioning datasets was not excluded, because the raw data was not available. Notably, we obtain an accurate fit between observed and simulated behavior if we rescale the behavioral traces by a factor of 1.3, which corresponds to excluding a hypothetical population of 23% non-responders (Figure 4.9 A). Hence, the

model $h_1$ which was calibrated on absolute conditioning data provides an explanation for this differential conditioning experiment: The model suggests a biologically plausible high degree of generalization between rewarded and unrewarded stimuli at low odor concentrations, and no operating mechanism for stimulus differentiation during training. It should be noted that the original study by Fernandez et al. (2009) found a significant degree of discrimination in a subsequent test phase, which may be explained by a retarded discrimination mechanism.

**Negative and positive patterning** We here consider two datasets on negative patterning. In the first dataset only one of the two stimuli of the unrewarded compound stimulus AB was presented with reinforcement (Figure 4.6 J). In the second dataset both stimuli A and B were presented with reinforcement (Figure 4.6 K). Treating the overlap between stimuli A and B as a free parameter (Equation 4.15, $overlap = 0.5$), we can readily explain the observed learning dynamics in the first dataset by the model $h_1$. Interestingly, the model suggests that the observed discrimination between rewarded (A) and unrewarded (AB) stimuli in this protocol results from only moderate generalization, however not from an active discrimination mechanism.

For the second negative patterning dataset, the model can explain a constant degree of discrimination between rewarded (A,B) and unrewarded stimuli (AB), however it is not compatible with the decreasing CR probabilities towards the end of the conditioning phase (Figure 4.6 K, $T > 12$). We recognize that our choice of MB encoding scheme creates a large substrate for inhibitory learning. For $overlap = 0.5$, 174 KCs are active exclusively to the compound stimulus AB (Figure 4.3 D). Adding this mechanism to the basic model $h_1$ allows to capture the decrease in CR probabilities to the unrewarded compound stimulus (Figure 4.9 B). As an alternative, we also found that a gradual decrease of AL activity to all stimuli increases the compatibility of simulated and observed behavior for this protocol (traces not shown). This decrease of AL activity may stem from a failed attempt to categorize between rewarded and unrewarded stimuli in the antennal lobe, which is presumably impossible during negative patterning. The implementation of biologically inspired categorization mechanisms in the AL may be the focus of future work (see Discussion).

The basic model $h_1$ can explain a constant degree of discrimination between rewarded (AB) and unrewarded (A,B) stimuli observed in positive patterning (Figure 4.6 L). Again, this is not a result of an active mechanism that learns to discriminate between non-rewarded elemental stimuli and rewarded compound stimuli, but simply results from the amount of generalization in our network. We also noticed that discrimination in positive patterning was high over a wide range of possible stimulus overlaps, and hence "easier to learn" than negative patterning, which resembles experimental observations in the fruit fly (Young et al., 2011). A further analysis of learning during negative and positive patterning would require the raw data sets, because only then the behavior of responding animals to unrewarded stimuli can be assessed. Unfortunately the raw data was not available at the time of this study.

## 4.4 Discussion

The goal of this study was to capture and explain behavior as observed in a variety of different conditioning protocols in the honeybee by a set of computational functions in a neuronal network model. After calibrating this model on behavioral data from absolute conditioning, we explored how well our model generalized to non-absolute conditioning protocols, such as differential conditioning and negative or positive patterning. We found that the basic model could readily account for behavior in non-absolute conditioning paradigms, if the degree of odorant similarity between the two employed toy stimuli was treated as a free parameter. We also explored two possible model extensions that were necessary to account for a gradual improvement of discrimination in differential conditioning, namely stimulus decorrelation in the antennal lobe, and inhibitory learning at the Mushroom body output. As we showed, both mechanisms were compatible with behavior in differential conditioning. Due to the increased model complexity, this result is little surprising, however our approach enables us to further test the eligibility of both extended models on any other conditioning protocol. For example, the second model extension ($h_1$+MB inhibitory learning) also correctly predicts observations in other protocols, such as latent inhibition (Chandra et al., 2010) or extinction (Stollhoff et al., 2005), whereas the first model extension ($h_1$+AL learning) cannot account for these findings.

The process of finding the best general model of associative learning during PER conditioning in honeybees naturally requires a more complete set of behavioral data first. Several important training conditions are missing in the current study, such as blocking (Guerrieri et al., 2005; Smith and Cobey, 1994; Gerber and Ullrich, 1999), reversal learning (Hadar and Menzel, 2010), backward conditioning (Hellstern et al., 1998), or conditioned inhibition. Quantitative predictions can be readily computed for these additional training conditions and compared to the actual behavioral observations. For example, for a blocking protocol, in which prior training to one stimulus (A+) is followed by training to a compound stimulus (AB+), our model would predict facilitated learning of the compound stimulus (traces not shown). It should be noted that this behavior is a consequence of our learning rule which does not make use of a total value signal. In contrast, Rescorla-Wagner-like learning rules in which learning is driven by the difference between the reward and the value of all present stimuli predict an impaired learning process for the compound stimulus. In our model, value could be implemented by post-synaptic activity in a three factor learning rule (Young et al., 2011).

In our study we employed group-averaged CR probabilities over trails as target traces for the simulations (Equation 4.15). For some of the protocols, in which the raw data was available, we excluded the proportion of non-responders from the raw data prior to computing these targets. For cases in which the raw data was not available, we exemplified that rescaling the target traces may substitute the exclusion of non-responders at the level of raw data (Figure 4.9 A). These procedures removed a large proportion of learning heterogeneity from

each dataset. However, several subtle differences in learning performance between individuals are not captured in these traces, and consequently missed by the model. For example, in absolute conditioning, animals responding early ($t_{firstCR} = 2$) show higher subsequent CR probabilities than animals responding later ($t_{firstCR} = 3, 4$) (Pamir et al. submitted). To give a second example, in differential conditioning, animals responding early also show higher amounts of initial generalization to the unrewarded odor, than animals responding later (Pamir et al. submitted). (This trend was also observed in another unpublished dataset which was recorded by Neloy Kumar Chakraborty). Our current model does not address these subtle differences in individual behavior among responding animals. Alternative to the employed error function (Equation 4.15), one may also compute an error on the basis of the raw data, which then allows formulating model hypotheses that can capture inter-individual differences in learning performance. However, this approach requires the raw data for all simulated conditioning protocols, which was not available at the time of this study. It should be noted that when comparing behavioral data from different studies our approach suffers from a typical high degree of variability across different PER conditioning experiments (Matsumoto et al., 2012; Frost et al., 2012). Preferentially, our study would build on a collection of data which was recorded in parallel, as in similar theoretical work in the fruit fly (Young et al., 2011; Wessnitzer et al., 2012). This would then also enable a more thorough analysis of the effect of stimulus parameters such as odorant similarity or relative size on learning under different conditions. Future experiments in this direction are under way.

Smith et al. (2012) identified an important computational problem in the in honeybee brain: "How can a neuropil (e.g. the MB) that receives a reinforcement (teaching) signal encode a memory for a pattern of input when that pattern may be changing as a result of the same reinforcement signal operating at an earlier stage of processing (e.g. the AL)?" In one of our model extensions ($h_1$+AL learning) we implemented a simple scheme for such a distributed plasticity mechanism, however at a high level of abstraction. More detailed hypotheses on reward dependent preprocessing in the antennal lobe exist (Rath et al., 2011), which may be implemented in our model. Theoretical studies also described non-reward dependent pre-processing (Linster et al., 2005; Schmuker et al., 2011) and plasticity in the antennal lobe (Assisi et al., 2012). Studying plasticity mechanisms that act on biologically realistic spatio-temporal odor traces in the antennal lobe (Fernandez et al., 2009) will ultimately require a spiking implementation of the current model.

## 4.5 Tables and table captions

| Conditioning protocol | Key finding | Reference |
|---|---|---|
| Absolute conditioning | Individuals rapidly acquire a stable CR | Pamir et al. 2011, submitted |
| Absolute conditioning with GRS pretest | Honeybees with high GRS show optimal learning, honeybees with low GRS show poor learning | Scheiner et al., 2001b, 2004; Pamir et al., submitted |
| Massed training trials | Short inter-trial-intervals reduce acquisition | Menzel, 2001; Pamir et al., submitted |
| Trace conditioning | Increasing the CS-US gap results in poorer acquisition | Szyszka et al., 2011; Pamir et al., submitted |
| Latent inhibtion | Pre-exposure to the unrewarded CS results in retarded acquisition | Chandra et al., 2010 |
| Differential conditioning | Response probabilities to the CS+ are unaffected, response probabilities to the CS- decrease over trials | Pamir et al., submitted |
| Differential conditioning with low concentrations | Poor discrimination between CS+ and CS- during training | Fernandez et al., 2009 |
| Differential conditioning to mixtures at different ratios | Poor discrimination between CS+ (9:1) and CS- (1:9) during training | Fernandez et al., 2009 |
| Negative patterning | Honeybees show differential responses, but no increase in discrimination | Unbublished data by Paul Szyszka |
| Negative patterning | Honeybees show differential responses, but no increase in discrimination | Deisig et al., 2001 |
| Positive patterning | Honeybees show differential responses, but no increase in discrimination | Deisig et al., 2001 |

**Table 4.1** Key findings on behavioral plasticity during classical conditioning of the proboscis extension response.

| Protocol | ITI (min) | Trial sequence | Odors | Odor conc. | Odor onsets (s) | Odor dur. (s) | Scrose conc. | Reward onset | Reward dur. |
|---|---|---|---|---|---|---|---|---|---|
| absolute | 30 | 6×A+ | 6- or 7-pentadecene | | 0 | 5 | 30% | 3 | 4 |
| absolute GRS test | 5 | 10×A+ | citral | $1\mu l/10ml$ | 0 | 4 | 30% | 3 | 2 |
| massed | 0.5 | 12×A+ | 2-hexanol | | 0 | 4 | 1.25 M | 3 | 3 |
| delay | 10 | 6×A+ | 1-nonanol | 1:100 mineral oil | 0 | 6 | 1 M | 5 | 3 |
| trace | 10 | 6×A+ | 1-nonanol | 1:100 mineral oil | 0 | 0.5 | 1 M | 1,5,10 | 3 |
| differential | 15 | 6×(A+,B-) | 1-hexanal,1-octanol | | 0 | 5 | 30% | 3 | 4 |
| differential low conc. | 6 | A+,B-,B-,A+,B-,A+,A+,B-,A+,B-,A+,B- | 1-hexanol, 2-octanone | 0.02 M | 0 | 4 | 2 M | 3 | 3 |
| differential mixture | 6 | A+,B-,B-,A+,B-,A+,A+,B-,A+,B-,A+,B-,B-,A+,B-,A+ | 9:1, 1:9, (1-hexanol,2-octanone) | 0.02 M | 0 | 4 | 2 M | 3 | 3 |
| negative patterning only A+ | 10 | 5×(A+,AB-) | | | 0 | 6 | | 5 | 3 |
| negative patterning | 8 | 6×(A+,B+,AB-,AB-) | linalool, limonene, 1-hexanol,2-octanol | pure | 0 | 6 | 1.25 M | 3 | 3 |
| positive patterning | 8 | 6×(AB+,AB+,A-,B-) | linalool, limonene, 1-hexanol,2-octanol | pure | 0 | 6 | 1.25 | 3 | 3 |

**Table 4.2** Summary of parameters of the simulated classical conditioning protocols. Abbreviations: inter-trial interval (ITI), seconds (s), molar (M);

| Parameter | Value | Description |
|---|---|---|
| $N_{PN}$ | 49 | Number of projection neurons |
| $N_{KC}$ | 5000 | Number of Kenyon cells |
| $N_{EN}$ | 1 | Number of Extrinsic neurons |
| $p_{PN \to KC}$ | 0.204 | Connection probability between PNs and KCs |
| $\theta_{KC}$ | 1.5 | Activity threshold of KCs |
| $\alpha_e$ | $0.05/s$ | Growth rate of KC eligibility trace |
| $\tau_e$ | $6.9\ s$ | Decay constant of KC eligibility trace |
| $\alpha_{oct}$ | $0.03/s$ | Growth rate of octopamine signal |
| $\tau_{oct}$ | $0.5\ s$ | Decay constant of octopamine signal |
| $\tau_{STM}$ | $36\ s$ | Time constant of short-term consolidation |
| $a_{EN}$ | 0.006 | Scaling factor of EN activity |
| $EN_{max}$ | 0.95 | Maximal EN activity |
| $S$ | 1.0 | Sensitivity to sucrose |
| $\alpha_w$ | $0.68/trial$ | Learning rate of KC weights |

**Table 4.3** Summary and description of parameters of the basic model $h_1$.

## 4.6 Figures and figure captions

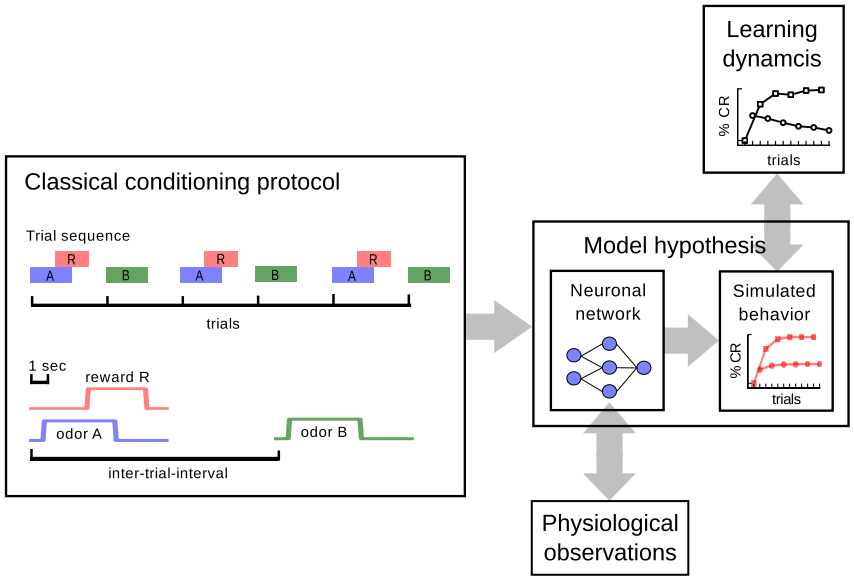


**Figure 4.1** A neuronal network approach to cognitive neuroethology in the honeybee. Cognitive learning faculties of honeybees have been assessed in a multitude of classical conditioning protocols. The employed conditioning protocols (e.g. differential conditioning as shown) define the sensory input to the neuronal network model at each trial. The sensory input is transformed to motor output via a series of computations, that mimic the computations in the honeybee brain. The motor output from the network is compared to the experimentally observed behavior during classical conditioning. Physiological observations on olfactory stimulus processing and neuronal plasticity in the honeybee brain constrain the computational functions performed by the neuronal network model.

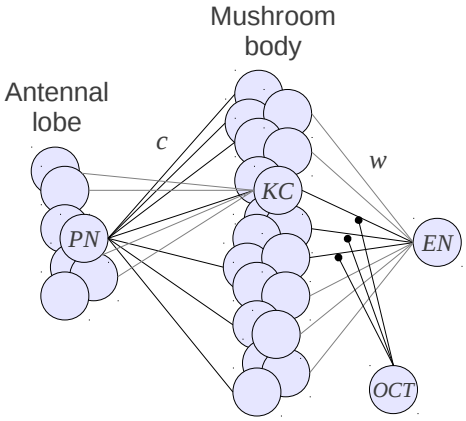


**Figure 4.2** Network geometry of the basic model $h_1$. On each trial, olfactory stimuli define a spatial activation pattern in the antennal lobe, which is represented by a 2-dimensional grid of size $7 \times 7$. 49 projection neurons (PN) project glomerular activation patterns to the Mushroom body Kenyon cells (KC) via a connection matrix $c$. Each KC receives input from exactly 10 PNs. All 5000 Mushroom body KCs connect to a single extrinsic neuron (EN) via a synaptic weight vector $w$. An octopaminergic neuron (OCT) provides a reinforcement signal to the KC synaptic weights. The activation of the extrinsic neuron equals the probability for a proboscis extension.

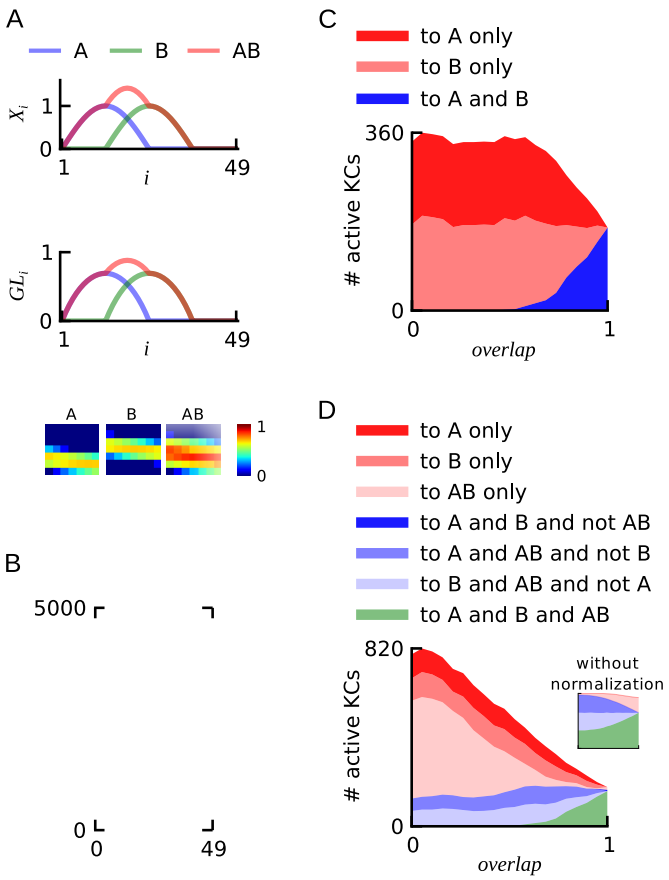


**Figure 4.3** The effect of stimulus overlap on Mushroom body encoding. **A** Stimuli set employed for the simulations. The glomerular activation patterns for odor A and B are defined by the positive part of a sinus function (Equation 4.1, 4.2). The overlap in glomerular activation between odor A and B is adjusted by shifting odor B to the right side, the depicted case has an overlap of 0.5. The compound stimulus AB is defined by adding the activation patterns of the two stimuli. Activation patterns of all three stimuli are transformed by a logarithmic transfer function (lower panel) (Equation 4.3). Activation patterns can be displayed on a $7 \times 7$ dimensional grid, that mimics the spatial activation patterns in the antennal lobe. **B** Pseudo-randomized connectivity matrix $c$ between 49 projection neurons and 5000 Kenyon cells. A white entry denotes a connection strength of unity, and a black entry denotes a connection strength of zero. The connection probability equals $p_{PN \to KC} = 0.204$. Each Kenyon cell receives input from exactly 10 projection neurons, and each projection neuron connects to 1020 Kenyon cells. **C** MB encoding scheme for the elemental stimuli A and B as a function of stimulus overlap between A and B. **D** MB encoding scheme for the elemental stimuli A and B, and the compound stimulus AB as a function of stimulus overlap between A and B. The small pictogram shows the encoding scheme without normalizatin in the antennal lobe. No Kenyon cells are exclusively active to the elemental stimuli A and B. Values learned for the elemental stimuli are always retrieved when the compound AB is presented.

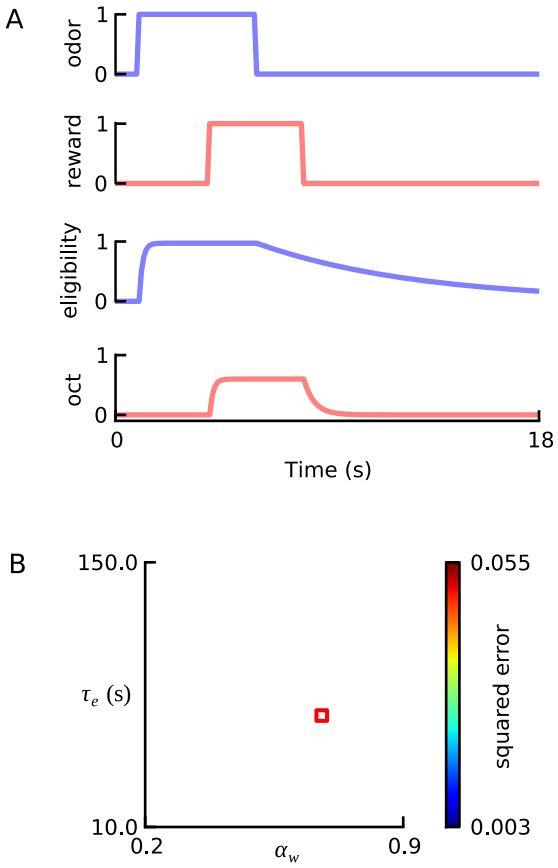


**Figure 4.4 A** Example traces for an odor (first trace) and a reward stimulus (second trace). The third trace shows the eligibility trace for the change in synaptic weight of a KC activated by odor A. The forth trace shows the trace of the octopamine concentration at the synapse signaling reinforcement. The weight change depends on the temporal overlap between the eligibility trace and the reinforcement signal. **B** Outcome of the parameter search for the learning rate $\alpha_w$ and the decay constant $\tau_{elig}$ of the eligibility trace. Shown are the summed normalized squared errors (4.16) estimated on five absolute classical conditioning protocols (absolute, delay, trace 1s, trace 5s, trace 10s). The red square marks the parameter combination of the basic model $h_1$, which minimizes the error ($\tau_{elig} = 6.9s$ and $\alpha_w = 0.716$).

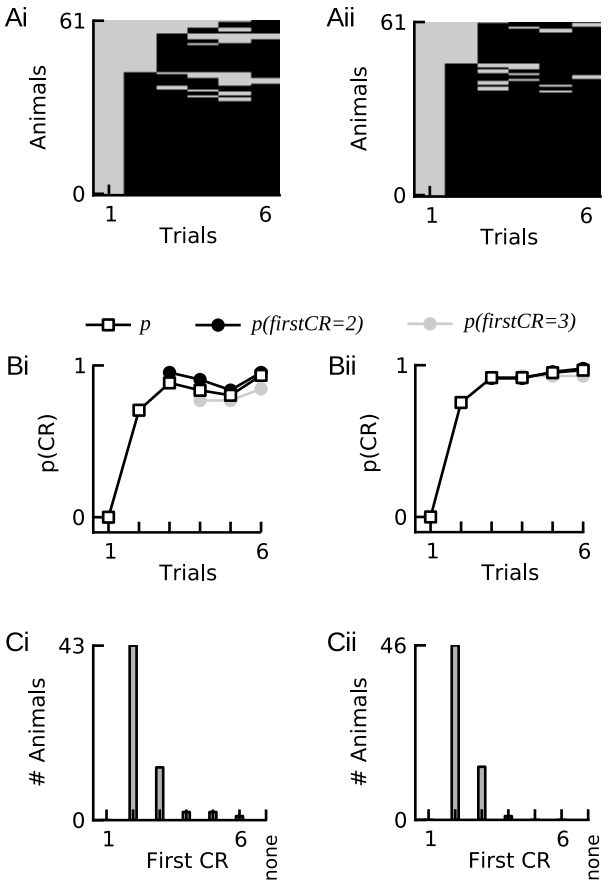


**Figure 4.5** The model $h_1$ is compatible with the behavioral features of individual learning during absolute classical conditioning, as described in Pamir et al. (submitted). The left column shows experimental data and the right column shows simulated data for the protocol absolute. **Ai, Aii** Binary CR matrix. A gray entry denotes no CR, and a black entry denotes a CR. **Bi, Bii** Group-average response probabilities ($p$) and conditional probabilities for animals showing their first CR on the second ($p(t_{firstCR}$=2)), or third trial ($p(t_{firstCR}$=3)) respectively. **Ci,Cii** Histograms of first CR latencies in trial-time.

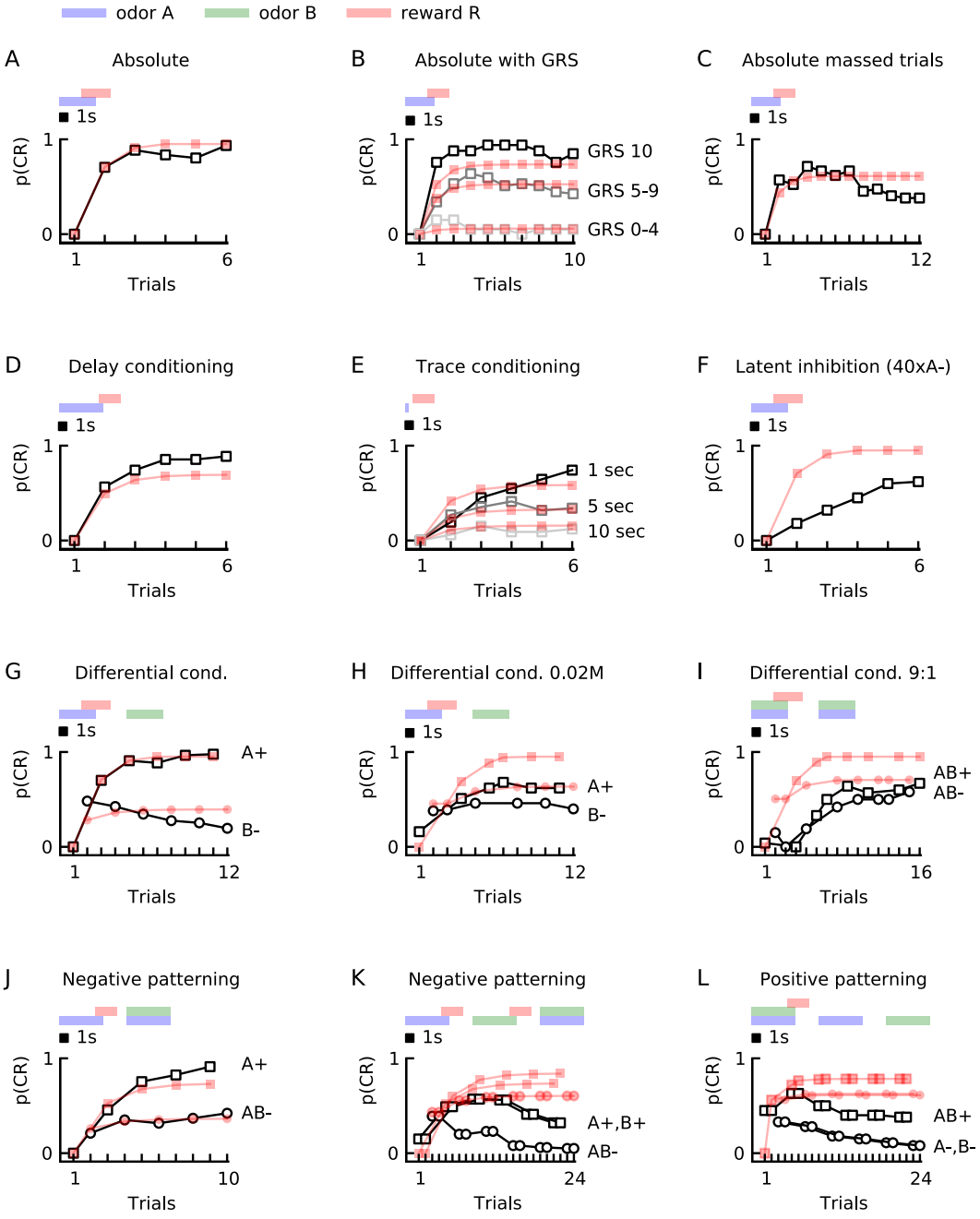


**Figure 4.6** Observed and simulated dynamics of associative learning in several different classical conditioning protocols. The black curves show the experimentally observed CR probabilities, and the red curves show the simulated CR probabilities of the basic model $h_1$. For the non-absolute classical conditioning protocols (G-L), the overlap between the two odors A and B was treated as a free parameter. Stimulus onsets and durations in each trial are depicted by colored pictograms. [*Caption continues on next page*]

[*Continuation of Figure 6 caption*] **A** A typical absolute conditioning protocol. The colored pictogram shows the first trial. Behavioral raw data by Neloy Kumar Chakraborty. **B** Absolute conditioning protocol with sucrose sensitivity pretest. The gustatory responsiveness score (GRS) of individuals was measured before the conditioning phase. Behavioral raw data by Ricarda Scheiner (Scheiner et al., 2001b). The chosen learning rule (Eq. 4.14) captures the correlation between GRS and learning performance. **C** Absolute conditioning protocol with short inter-trial intervals (massed conditioning trials). Behavioral raw data by Randolf Menzel (Menzel, 2001). A triggered change in synaptic weight requires a certain time to consolidate. Implementing this mechanism (4.13) captures the dependence of learning performance on the inter-trial-interval. **D, E** Delay conditioning and trace conditioning at three different inter-stimulus intervals (1, 5, 10 seconds). Behavioral raw data by Paul Szyszka (Szyszka et al., 2011). By implementing an eligibility trace in the KCs, the basic model captures the dependence of the learning performance on the inter-stimulus interval. **F** Latent inhibition protocol (pre-conditioning phase with 40 unrewarded trials (A-) is not shown). The behavioral traces were transcribed manually from Figure 3B in Chandra et al. (2010). The basic model is not compatible with the observation of a retarded acquisition after 40 unrewarded conditioned stimuli, as described in Chandra et al. (2010). **G** Differential conditioning protocol at high odor concentrations. The colored pictogram shows the first two trials. Behavioral raw data by Neloy Kumar Chakraborty. The basic model captures the learning dynamics of the rewarded stimulus A+, however it does not capture the gradual increase in discrimination between rewarded (A+) and unrewarded stimuli (B-) which is typically observed in differential conditioning. **H, I** Differential conditioning with odors at low concentrations and mixture stimuli at low concentrations. The behavioral traces were transcribed manually from Figure 1 A and B in Fernandez et al. (2009). The basic model captures the poor degree of discrimination during the training phase for the two conditions. **J** Negative patterning protocol in which only one of the two stimuli (A+) of the unrewarded compound (AB-) was rewarded. Behavioral raw data by Paul Szyszka (unpublished). The basic model captures the learning dynamics observed under this protocol. **K, L** Negative and positive patterning. The colored pictograms show the first three trials. The behavioral traces were transcribed manually from Figure 1B and 2B in Deisig et al. (2001). The basic model captures stronger responses to the rewarded than to the unrewarded stimuli in negative and positive patterning. However, the model does not account for a gradual decline of response probabilities observed in the two protocols.

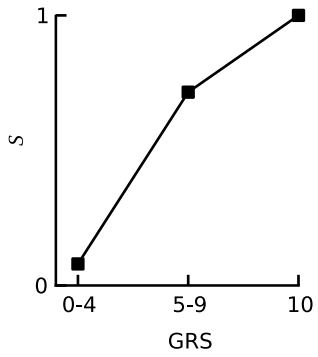


**Figure 4.7** Mapping between experimentally assessed gustatory response scores (GRS) and reward evaluation within the neuronal network model $h_1$.

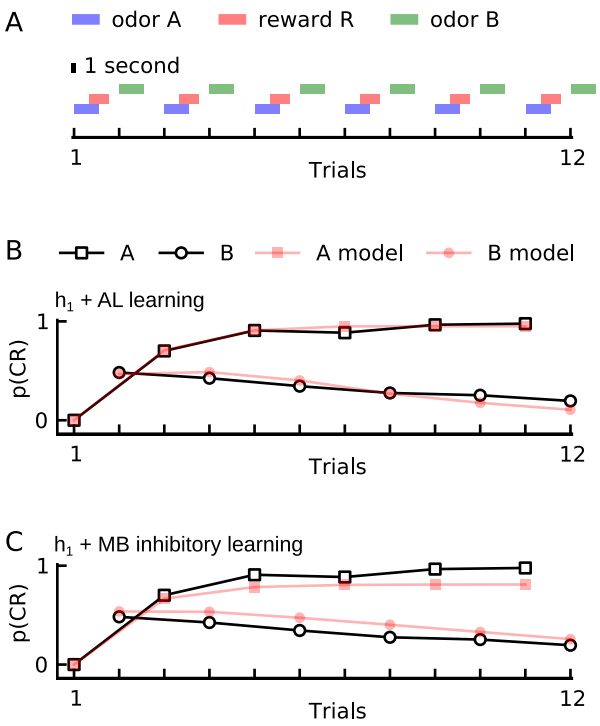


**Figure 4.8** Two different hypotheses can explain the gradual increase in discrimination between rewarded and unrewarded stimuli observed during differential conditioning. **A** Employed conditioning protocol for experiments and simulations. **B** The basic model $h_1$ captures the behavioral data if one assumes an additional learning mechanism in the antennal lobe, which decreases the overlap between rewarded and unrewarded stimuli. **C** As an alternative, the basic model plus an inhibitory learning mechanism in the Mushroom body KCs also captures the behavioral data.

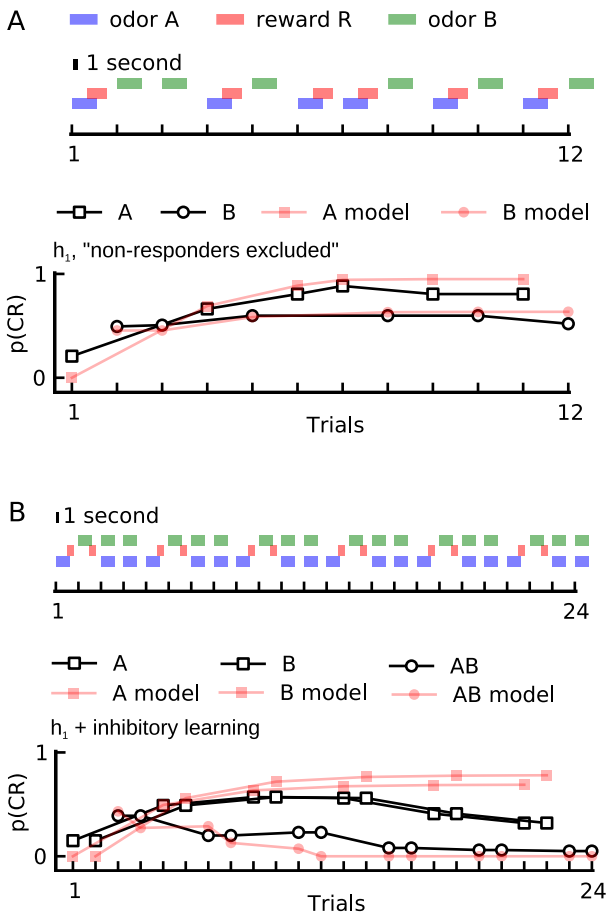


**Figure 4.9** Additional assumptions that increase the compatibility of the basic model $h_1$ with behavioral data. **A** The basic model captures the behavior observed in differential conditioning with odors at low concentrations if one excludes a hypothetical proportion of non-responders from the dataset. Since the raw data was not available, we rescaled the behavioral traces by a factor of 1.3, which corresponds to excluding a hypothetical proportion of 23% non-responders. Data origin as in Figure 4.6 H. **B** The gradual increase in discrimination between rewarded (A+,B+) and unrewarded stimuli (AB-) observed in negative patterning can be captured by adding an inhibitory learning mechanism to the basic model. Data origin as in Figure 4.6 K.

# Chapter 5

# General discussion

## 5.1 A new perspective on behavioral data and learning in the honeybee

It has long been recognized that honeybees rapidly learn the association between the conditioned and unconditioned stimulus during classical conditioning of the proboscis extension response. Bitterman et al. (1983) found that about 30% to 80% of the honeybees in each group of trained animals showed a conditioned response already on the second trial. In order to clarify the dynamics of this seemingly fast learning process, the authors of this study looked at group-average learning curves in different training groups, and suggested that the associative strength would probably require several, possibly eight, conditioning trials in order to reach its asymptotic value. In Chapters 2 and 3 of this thesis I presented an alternative view on the behavioral data in the honeybee. As was impressively proven for several learning paradigms in vertebrates, group-average learning curves or acquisition functions can hide the actual dynamics of learning in the individual (Gallistel et al., 2004). As I showed, the same applies to the honeybee. Analyzing first-order serial correlations in conditioned responses I showed that average behavior does not represent individual behavior during classical conditioning of the proboscis extension response (Chapter 2). Importantly, and as also noted by Scheiner et al. (2005), the high degree of inter-individual variability found in learning performance questions the predominant use of group-average memory retention scores in the experimental literature.

To provide an alternative to the group-average perspective on learning and memory in the honeybee, Chapter 3 presents a novel parametric description for associative learning. In a given dataset, individual learning can be characterized by three parameters: (1) The ability of the individual to learn the task, as indirectly measured by the percentage of non-responders, (2) the response latencies of individuals measured in trial time, (3) and the stability of subsequent responses in responding animals. Based on these parameters, Chapter 3 presented a detailed analysis of the dynamics of individual learning in a large collection of data on absolute

and non-absolute conditioning. From this analysis followed that individual behavior was characterized by a rapid and stable acquisition of the conditioned response. On average, 94% of all animals showing any response at all acquired the conditioned response within the first four conditioning trials. Subsequent responses in these animals were very stable, irrespective of when the conditioned response had been acquired during conditioning. In each data set, a stable proportion of animals (typically 10% to 30%) did not respond in any of the trials. Based on experimental work by Ricarda Scheiner (Scheiner et al., 1999, 2001a), I showed that non-responders were characterized by a poor sensitivity to sucrose.

Given these results, the group-average learning curve has to be reinterpreted as follows: (1) The gradual rise of the curve on the first few conditioning trials cannot be taken as behavioral evidence for a gradual increase in associative strength in individuals. The gradual rise simply reflects the trial-by-trial recruitment of honeybees with a stable response. (2) The asymptote of the group-average learning curve reflects the percentage of non-responders in each data set, as well as the actual stability of the conditioned response in responding animals. However, it does not reflect a performance asymptote of individual learning. Shifting the focus of analysis from the training phase to memory, in Chapters 2 and 3 I found that memory retention in individuals complies with a constant continuation of behavior expressed during the conditioning phase. Importantly, this means that group-average memory retention does not represent memory strength or stability in individuals, but mainly the percentage of animals that acquired a stable CR during the training phase.

This new perspective on learning and memory at the level of individuals called into question some long-held assumptions about the effect of training parameters on the induction of different memories in the honeybee. The current standard model distinguishes between single-trial and three-trial induced memories, which are characterized by different properties both at the behavioral and physiological level of analysis (Menzel, 1990; Müller, 2012; Menzel, 2012). At the group-average behavioral level, evidence for this memory model was provided by the observation of higher memory retention after three-trial conditioning than after single-trial conditioning when tested 24 hours after training (Menzel, 1990, Fig. 9.8). In an experimental collaboration with Paul Szyszka from the University of Konstanz I could show that memory retrieval at this time point did not depend on the number of conditioning trials per se. Animals that underwent single-trial conditioning showed equal levels of memory retention and discriminatory power as animals that underwent two-trial or four-trial conditioning, given that animals acquired the conditioned response during training (Chapter 3). Taken together, no behavioral evidence has been found in this thesis that more training results in a more stable or specific behavioral expression of memory in individuals. On the contrary, typically early responders were found to show slightly higher levels of response probabilities in subsequent trials than late responders (Chapter 3).

Regarding evidence from biochemistry (as reviewed in Müller, 2012), the role of training

intensity on molecular events in individual brains remains unclear. From the experimental literature, it is difficult to understand to what extent individual behavior has been factored into the biochemical analysis of brain tissues from honeybees that underwent classical conditioning. To give an example, Müller (2000) found a prolongated activation of protein kinase A after three-trial, but not after single-trial conditioning. The author employed fast-freezing techniques (Hildebrandt and Müller, 1995), which allowed capturing the molecular state of the brain directly after training. However, brain tissues were analyzed at the group-average level, probably without knowledge of individual behavior. One can speculate that a more informative or possibly different experimental result could be obtained by comparing molecular signatures of memory consolidation in individuals with a conditioned response history of 0(1) and 0111 (see Chapter 3). This would then elucidate the effect of training intensity on the induction of molecular processes in the actual individual brain, in contrast to the described effects on the "group-average brain". The same sort of speculative reasoning may be applicable to the current model of memory in the fruit fly (Davis 2011), which has been predominantly established at the group-average level. It seems that only recently studies in the fruit fly (Chabaud et al., 2010), and in the honeybee (Roussel et al., 2010; Rath et al., 2011; D'Albis et al., 2011; Mota and Giurfa, 2010) have started to incorporate individual behavior in their analysis of experimental data, and maybe this thesis provides further impulses in this direction.

## 5.2 What can be learned about learning by computational modeling?

Understanding how neuronal functioning in different areas of the insect brain gives rise to associative learning in the behaving animal is a complex puzzle to solve. As pointed out by two recent theoretical studies, solving this puzzle requires a computational framework, in which both behavioral and neurophysiological evidence can be integrated (Smith et al., 2012; Wessnitzer et al., 2012). For this, the establishment of a detailed understanding of the temporal dynamics of associative learning in individual animals seems to be a basic prerequisite. In this thesis, I tried to establish a refined account of behavioral plasticity as observed during classical conditioning of the proboscis extension response in the honeybee (Chapters 2, 3). In turn, this had direct consequences for the theoretical account of associative learning at all levels of model complexity (Chapters 2, 3, 4).

The Rescorla-Wagner model (Rescorla and Wagner, 1972), which was originally developed to conceptualize various behavioral phenomena observed in classical conditioning experiments, has to date provided an influential theory of associative learning (Dayan and Abbott, 2001; Gluck and Myers, 2001). Fitting this model to classical conditioning data results in a description of learning by two parameters: the learning rate, which is assumed to depend on the salience of the conditioned stimulus and the unconditioned stimulus, and the maximum asymptotic

value of associative strength supported by the unconditioned stimulus (Rescorla and Wagner, 1972). As I showed in Chapter 2 and 3 of this thesis, the informative value of this approach crucially depends on *how* the model is fit to a sample of data from identically treated animals. In the case of the honeybee, where the probabilistic expression of behavior in the individual does not follow the group-average performance (Pamir et al., 2011), fitting and interpreting this theory at the group-average level is erroneous. The Rescorla-Wagner theory, however, can be easily rescued by fitting the model at the level of individuals, for example by introducing some sort of learning heterogeneity in the formalism (Chapters 2, 3). By this approach, one can then compute and illustrate the learning heterogeneity in a sample of honeybees as well as its modulation under altered experimental conditions (Chapter 3). A final distinction between different alternative hypotheses on the actual type of heterogeneity was hard to establish based on the experimental data. Cross-validating different hypotheses against each other did not reveal a definite result, except that some sort of heterogeneity is absolutely required to correctly model the data, and that slow learning is not a plausible explanation for non-responders (Chapters 2, 3). In Chapter 3, I discussed a "final version" of an extended Rescorla-Wagner model which, considering all experimental evidence, presents the most likely working hypothesis so far.

As I showed in Chapter 2, the behavioral dynamics of learning in individual honeybees can also be well captured by a two-state hidden Markov model. However, in contrast to the Rescorla-Wagner model, in which learning is driven by prediction errors, the Markov model does not advise any biologically motivated mechanisms for learning. Rather it advises a dynamical system, defined by a set of constant transition and observation probabilities, which after the initiation of the experiment relaxes to its equilibrium state. It seems that this type of model is better suited to study the probabilistic expression of behavior within a constant or innate behavioral repertoire, as for example observed during chemotaxis in larval Drosophila (Gomez-Marin et al., 2011).

The modeling approach outlined so far (Chapters 2, 3) provided a better understanding of the behavioral dynamics of associative learning in individuals. However, it did not make contact with the experimental literature on neuronal functioning in the insect brain, which is the goal of building integrative models (Wessnitzer et al., 2012). In Chapter 4 of this thesis I presented a model hypothesis on olfactory information processing and associative learning in the honeybee brain circuitry. This model provides a framework for studying the effect of external stimulus parameters, internal stimulus encoding schemes and computational principles such as divisive normalization or associative learning rules on the dynamics of behavioral plasticity observed during various different learning tasks. As was already the case for the preceding model approach (Chapter 2, 3), a detailed analysis of behavior provided the basis for this. Computational research in the honeybee can therein draw from an ample collection of learning data (reviewed in Menzel et al., 2007; Sandoz, 2011) that was recorded

at the level of individuals over the full time course of the experiment. The same is not true in Drosophila. Due to differences in methodology, computational studies have to rely on singular group-average performance scores, typically assessed only after the end of the training phase (Young et al., 2011; Wessnitzer et al., 2012). This may be problematic for the computational approach: In learning tasks such as differential conditioning and positive or negative patterning, a final performance score does not reveal if individuals actually learned to discriminate rewarded from non-rewarded conditions. A significant performance score may as well be a sole product of moderate stimulus generalization within the circuitry, as also noticed by Wessnitzer et al. (2012). As an alternative, in the honeybee, the model constraints collected for Chapter 4 include explicit temporal dynamics during training, which provides a more informative behavioral fingerprint of neuronal functioning.

Admittedly, also in the honeybee, the goal of building integrative models suffers from several limitations and drawbacks. As pointed out in the general introduction of this thesis (Chapter 1), it is often difficult to establish a causal linkage between findings on neuronal plasticity in the brain and learning at the behavioral level (Gallistel and Matzel, 2012). This problem of disparate properties of neuronal and behavioral plasticity is also evident in the experimental literature in the honeybee. Plasticity is typically detected and quantified by comparing neuronal activity that was recorded before and after the conditioning phase, and hence the observed induction time of plasticity rather reflects experimental routines than biological kinetics. For example, changes in neuronal activity to conditioned and unconditioned stimuli were observed 24 hours after differential conditioning in the antennal lobe (Fernandez et al., 2009), and 3 hours after differential conditioning in the mushroom body (Strube-Bloss et al., 2011). It remains an open question how model constraints on neuronal plasticity during the actual training phase can be extracted from these findings.

To point to another problem, although behavioral plasticity has been recorded over trials in individuals, findings on learning are distributed across several decades of learning experiments. At the same time, the experimental routines have been further developed (Matsumoto et al., 2012; Frost et al., 2012), and possibly a direct comparison between different findings is inappropriate. Ideally, the approach detailed out in Chapter 4 would build on a large collection of data on absolute and non-absolute learning tasks with conserved stimuli sets. Chapter 4 may give advise for future combined experimental and theoretical work in this direction.

To put forward a final objection, computational approaches generally seem to lag behind experimental research. Following the working hypothesis in the field (Heisenberg, 2003; Gerber et al., 2004), computational studies have so far located associative learning in the mushroom body (Huerta and Nowotny, 2009; Wessnitzer et al., 2012). However, memory traces are distributed in time and space in the insect brain (Davis, 2011; Smith et al., 2012), and recently, the role of the mushroom body as the site of the memory trace has been challenged (Davis and Giurfa, 2012). What really can be learned about learning through theoretical approaches

remains the subject of future work, but possibly a renewed interest in behavioral plasticity may contribute to solving the puzzle.

# Bibliography


Arena P, Patané L, Stornanti V, Termini PS, Zäpf B, Strauss R (2012) Modeling the insect mushroom bodies: Application to a delayed match-to-sample task. *Neural networks : the official journal of the International Neural Network Society* .

Assisi C, Stopfer M, Bazhenov M (2012) Excitatory local interneurons enhance tuning of sensory information. *PLoS computational biology* 8:e1002563.

Behrends A, Scheiner R (2010) Learning at Old Age: A Study on Winter Bees. *Frontiers in behavioral neuroscience* 4:11.

Behrends A, Scheiner R (2012) Octopamine improves learning in newly emerged bees but not in old foragers. *The Journal of experimental biology* 215:1076–83.

Ben-Shahar Y, Robinson GE (2001) Satiation differentially affects performance in a learning assay by nurse and forager honey bees. *Journal of comparative physiology A* 187:891–9.

Bhandawat V, Olsen SR, Gouwens NW, Schlief ML, Wilson RI (2007) Sensory processing in the Drosophila antennal lobe increases reliability and separability of ensemble odor representations. *Nature neuroscience* 10:1474–82.

Biergans SD, Jones JC, Treiber N, Galizia CG, Szyszka P (2012) DNA methylation mediates the discriminatory power of associative long-term memory in honeybees. *PloS one* 7:e39349.

Bitterman ME, Menzel R, Fietz A, Schäfer S (1983) Classical conditioning of proboscis extension in honeybees (Apis mellifera). *Journal of comparative psychology* 97:107–19.

Brown S, Heathcote A (2003) QMLE: fast, robust, and efficient estimation of distribution functions based on quantiles. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc* 35:485–92.

Byrd R, Lu P, Nocedal J (1994) A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16:1190–1208.

Caporale N, Dan Y (2008) Spike timing-dependent plasticity: a Hebbian learning rule. *Annual review of neuroscience* 31:25–46.

Carandini M (2012) From circuits to behavior: a bridge too far? *Nature neuroscience* 15:507–9.

Cassenaer S, Laurent G (2012) Supplementary Information. Conditional modulation of spike-timing-dependent plasticity for olfactory learning. *Nature* 482:47–52.

Chabaud MA, Preat T, Kaiser L (2010) Behavioral characterization of individual olfactory memory retrieval in Drosophila melanogaster. *Frontiers in behavioral neuroscience* 4:192.

Chandra S, Smith BH (1998) An analysis of synthetic processing of odor mixtures in the honeybee (Apis mellifera). *The Journal of experimental biology* 201:3113–21.

Chandra SBC, Wright Ga, Smith BH (2010) Latent inhibition in the honey bee, Apis mellifera: Is it a unitary phenomenon? *Animal cognition* 13:805–15.

Cousineau D, Hélie S, Lefebvre C (2003) Testing curvatures of learning functions on individual trial and block average data. *Behavior research methods, instruments, & computers : a journal of the Psychonomic Society, Inc* 35:493–503.

D'Albis T, Haenicke J, Strube-Bloss MF, Schmuker M, Menzel R, Nawrot MP (2011) Learning-induced changes at the single neuron level predict behavioral performance in the honeybee. *Frontiers in neuroscience (event abstract)* .

Davis R (2011) Traces of Drosophila Memory. *Neuron* 70:8–19.

Davis RL, Giurfa M (2012) Mushroom-Body Memories: An Obituary Prematurely Written? *Current Biology* 22:R272–R275.

Dayan P, Abbott LF (2001) *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* The MIT Press.

Deisig N, Giurfa M, Lachnit H, Sandoz JC (2006) Neural representation of olfactory mixtures in the honeybee antennal lobe. *The European journal of neuroscience* 24:1161–74.

Deisig N, Lachnit H, Giurfa M (2002) The effect of similarity between elemental stimuli and compounds in olfactory patterning discriminations. *Learning & memory (Cold Spring Harbor, N.Y.)* 9:112–21.

Deisig N, Lachnit H, Giurfa M, Hellstern F (2001) Configural olfactory learning in honeybees: Negative and positive patterning discrimination. *Learning & Memory* pp. 70–78.

Deisig N, Lachnit H, Sandoz JC, Lober K, Giurfa M (2003) A modified version of the unique cue theory accounts for olfactory compound processing in honeybees. *Learning & memory (Cold Spring Harbor, N.Y.)* 10:199–208.

Dubnau J, Chiang AS, Tully T (2003) Neural substrates of memory: from synapse to system. *Journal of neurobiology* 54:238–53.

Dudai Y (2004) The neurobiology of consolidations, or, how stable is the engram? *Annual review of psychology* 55:51–86.

Dukas R (2008) Evolutionary biology of insect learning. *Annual review of entomology* 53:145–60.

Durstewitz D, Vittoz NM, Floresco SB, Seamans JK (2010) Abrupt transitions between prefrontal neural ensemble states accompany behavioral transitions during rule learning. *Neuron* 66:438–48.

Eisenhardt D (2006) Learning and memory formation in the honeybee (Apis mellifera) and its dependency on the cAMP-protein kinase A pathway. *Animal biology* 56:259–278.

Erber J, Kierzek S, Sander E (1998) Tactile learning in the honeybee. *Journal of comparative physiology A* pp. 737–744.

Estes WK (2002) Traps in the route to models of memory and decision. *Psychonomic bulletin & review* 9:3–25.

Faber T, Joerges J, Menzel R (1999) Associative learning modifies neural representations of odors in the insect brain. *Nature neuroscience* 2:74–8.

Feldman DE (2012) The spike-timing dependence of plasticity. *Neuron* 75:556–71.

Felsenberg J, Gehring KB, Antemann V, Eisenhardt D (2011) Behavioural pharmacology in classical conditioning of the proboscis extension response in honeybees (Apis mellifera). *Journal of visualized experiments : JoVE* .

Fernandez PC, Locatelli FF, Person-Rennell N, Deleo G, Smith BH (2009) Associative conditioning tunes transient dynamics of early olfactory processing. *The Journal of Neuroscience* 29:10191–10202.

Fiala A, Müller U, Menzel R (1999) Reversible Downregulation of Protein Kinase A during Olfactory Learning Using Antisense Technique Impairs Long-Term Memory Formation in the Honeybee, Apis mellifera. *The Journal of Neuroscience* 19:10125–10134.

Friedrich A, Thomas U, Müller U (2004) Learning at different satiation levels reveals parallel functions for the cAMP-protein kinase A cascade in formation of long-term memory. *The Journal of Neuroscience* 24:4460–4468.

Frost EH, Shutler D, Hillier NK (2012) The proboscis extension reflex to evaluate learning and memory in honeybees (Apis mellifera): some caveats. *Naturwissenschaften* pp. 677–686.

Galizia CG, Sachse S, Rappert A, Menzel R (1999) The glomerular code for odor representation is species specific in the honeybee Apis mellifera. *Nature neuroscience* 2:473–8.

Gallistel CR, Matzel LD (2012) The Neuroscience of Learning: Beyond the Hebbian Synapse. *Annual review of psychology* .

Gallistel CR, Fairhurst S, Balsam P (2004) The learning curve: implications of a quantitative analysis. *PNAS* 101:13124–31.

Gerber B, Ullrich J (1999) No evidence for olfactory blocking in honeybee classical conditioning. *The Journal of experimental biology* 202 (Pt 13:1839–54.

Gerber B, Wüstenberg D, Schütz A, Menzel R (1998) Temporal determinants of olfactory long-term retention in honeybee classical conditioning: nonmonotonous effects of the training trial interval. *Neurobiology of learning and memory* 69:71–8.

Gerber B, Tanimoto H, Heisenberg M (2004) An engram found? Evaluating the evidence from fruit flies. *Current opinion in neurobiology* 14:737–44.

Giovanni Galizia C (2011) *Honeybee Neurobiology and Behavior* Springer.

Giurfa M (2003) Cognitive neuroethology: dissecting non-elemental learning in a honeybee brain. *Current Opinion in Neurobiology* 13:726–735.

Giurfa M (2007) Behavioral and neural analysis of associative learning in the honeybee: a taste from the magic well. *Journal of comparative physiology A* 193:801–24.

Giurfa M, Sandoz JC (2012) Invertebrate learning and memory: Fifty years of olfactory conditioning of the proboscis extension response in honeybees. *Learning & Memory* 19:54–66.

Gluck MA, Myers CE (2001) *Gateway to Memory: An Introduction to Neural Network Modeling of the Hippocampus and Learning (Issues in Clinical and Cognitive Neuropsychology)* A Bradford Book.

Gomez-Marin A, Stephens GJ, Louis M (2011) Active sampling and decision making in Drosophila chemotaxis. *Nature communications* 2:441.

Guerrieri F, Schubert M, Sandoz JC, Giurfa M (2005) Perceptual and neural olfactory similarity in honeybees. *PLoS biology* 3:e60.

Hadar R, Menzel R (2010) Memory formation in reversal learning of the honeybee. *Frontiers in behavioral neuroscience* 4:186.

Hammer M (1993) An identified neuron mediates the unconditioned stimulus in associative olfactory learning in honeybees. *Nature* 366:59–63.

Hanson S, Killeen P (1981) Measurement and modeling of behavior under fixed-interval schedules of reinforcement. *J. Exp. Psychol.: Anim. Behav. Process.* 7:129–39.

Hebb DO (1949) *TheOrganization of Behavior: A Neuropsychological Theory.* New York: Wiley.

Heisenberg M (2003) Mushroom body memoir: from maps to models. *Nature reviews. Neuroscience* 4:266–75.

Hellstern F, Malaka R, Hammer M (1998) Backward inhibitory learning in honeybees: a behavioral analysis of reinforcement processing. *Learning & Memory* 4:429–444.

Hildebrandt H, Müller U (1995) PKA activity in the antennal lobe of honeybees is regulated by chemosensory stimulation in vivo. *Brain research* 679:281–288.

Himmelreich S, Grünewald B (2012) Cellular physiology of olfactory learning in the honeybee brain. *Apidologie* .

Huerta R (2013) Learning pattern recognition and decision making in the insect brain. *Physics* 119:101–119.

Huerta R, Nowotny T (2009) Fast and robust learning by reinforcement signals: explorations in the insect brain. *Neural computation* 21:2123–51.

Izhikevich EM (2007) Solving the distal reward problem through linkage of STDP and dopamine signaling. *Cerebral cortex (New York, N.Y. : 1991)* 17:2443–52.

Kazama H, Wilson RI (2008) Homeostatic matching and nonlinear amplification at identified central synapses. *Neuron* 58:401–13.

Krechevsky I (1932) "Hypotheses" in rats. *Psychol Rev* 39:516 –532.

Kreissl S, Eichmüller S, Bicker G, Rapus J, Eckert M (1994) Octopamine-like immunoreactivity in the brain and subesophageal ganglion of the honeybee. *The Journal of comparative neurology* 348:583–95.

Liang ZS, Nguyen T, Mattila HR, Rodriguez-Zas SL, Seeley TD, Robinson GE (2012) Molecular Determinants of Scouting Behavior in Honey Bees. *Science* 335:1225–1228.

Linster C, Sachse S, Galizia CG (2005) Computational modeling suggests that response properties rather than spatial position determine connectivity between olfactory glomeruli. *Journal of neurophysiology* 93:3410–7.

Locatelli FF, Fernandez PC, Villareal F, Muezzinoglu K, Huerta R, Galizia CG, Smith BH (2012) Nonassociative plasticity alters competitive interactions among mixture components in early olfactory processing. *European Journal of Neuroscience* pp. n/a–n/a.

Lockett GA, Helliwell P, Maleszka R (2010) Involvement of DNA methylation in memory processing in the honey bee. *Neuroreport* 21:812–6.

Matsumoto Y, Menzel R, Sandoz Jc, Giurfa M (2012) Revisiting olfactory classical conditioning of the proboscis extension response in honey bees: A step toward standardized procedures. *Journal of Neuroscience Methods* pp. 1–9.

Menzel R (1990) Learning, memory and "cognition" in honeybees In Kesner R, Olton D, editors, *Neurobiology of comparative cognition*, chapter 9, pp. 237–292. Hillsdale, N.J.

Menzel R (1999) Memory dynamics in the honeybee. *Journal of comparative physiology A* 185:323–340.

Menzel R (2001) Searching for the memory trace in a mini-brain, the honeybee. *Learning & Memory* 8:53–62.

Menzel R (2012) The honeybee as a model for understanding the basis of cognition. *Nature Reviews Neuroscience* 13:758–768.

Menzel R, Brembs B, Giurfa M (2007) Cognition in invertebrates . In *Evolution of nervous systems in invertebrates*, pp. 403–422. Academic Press, Oxford, UK.

Menzel R, Manz G, Menzel R, Greggers U (2001) Massed and Spaced Learning in Honeybees : The Role of CS , US , the Intertrial Interval , and the Test Interval. *Learning & Memory* 8:189–208.

Mota T, Giurfa M (2010) Multiple reversal olfactory learning in honeybees. *Frontiers in behavioral neuroscience* 4.

Müller U (2000) Prolonged activation of cAMP-dependent protein kinase during conditioning induces long-term memory in honeybees. *Neuron* 27:159–68.

Müller U (2012) The molecular signalling processes underlying olfactory learning and memory formation in honeybees. *Apidologie* .

Page RE, Erber J, Fondrk MK (1998) The effect of genotype on response thresholds to sucrose and foraging behavior of honey bees (Apis mellifera L.). *Journal of comparative physiology. A* 182:489–500.

Pamir E, Chakroborty NK, Stollhoff N, Gehring KB, Antemann V, Morgenstern L, Felsenberg J, Eisenhardt D, Menzel R, Nawrot MP (2011) Average group behavior does not represent individual behavior in classical conditioning of the honeybee. *Learning & Memory* 18:733–41.

Pelz C, Gerber B, Menzel R (1997) Odorant intensity as a determinant for olfactory conditioning in honeybees: roles in discrimination, overshadowing and memory consolidation. *The Journal of experimental biology* 200:837–47.

Perez-Orive J, Mazor O, Turner GC, Cassenaer S, Wilson RI, Laurent G (2002) Oscillations and sparsening of odor representations in the mushroom body. *Science (New York, N.Y.)* 297:359–65.

Quinn WG, Harris WA, Benzer S (1974) Conditioned behavior in Drosophila melanogaster. *PNAS* 71:708–12.

Rath L, Giovanni Galizia C, Szyszka P (2011) Multiple memory traces after associative learning in the honey bee antennal lobe. *The European journal of neuroscience* 34:352–60.

Rehder V (1987) Quantification of the honeybee's proboscis reflex by electromyographic recordings. *Journal of Insect Physiology* 33:501–507.

Rescorla R, Wagner A (1972) A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement In *Classical Conditioning II: current research and theory.*, pp. 64 – 99.

Restle F (1965) Significance of all-or-none learning. *Psychol Bull* 64:313 – 325.

Roussel E, Sandoz JC, Giurfa M (2010) Searching for learning-dependent changes in the antennal lobe: simultaneous recording of neural activity and aversive olfactory learning in honeybees. *Frontiers in behavioral neuroscience* 4.

Rudy JW, Sutherland RJ (1995) Configural association theory and the hippocampal formation: an appraisal and reconfiguration. *Hippocampus* 5:375–89.

Rudy JW, Sutherland RJ (1992) Configural and Elemental Associations and the Memory Coherence Problem. *Journal of Cognitive Neuroscience* 4:208–216.

Sandoz J, Roger B, Pham-Delègue MH (1995) Olfactory learning and memory in the honeybee: comparison of different classical conditioning procedures of the proboscis extension response. *Comptes rendus de lAcademie des sciences Serie III Sciences de la vie* 318:749–755.

Sandoz JC (2011) Behavioral and Neurophysiological Study of Olfactory Perception and Learning in Honeybees. *Frontiers in Systems Neuroscience* 5:98.

Scheiner R, Barnert M, Erber J (2003) Variation in water and sucrose responsiveness during the foraging season affects proboscis extension learning in honey bees. *Apidologie* 34:67–72.

Scheiner R, Erber J, Page RE (1999) Tactile learning and the individual evaluation of the reward in honey bees (Apis mellifera L.). *Journal of comparative physiology A* 185:1–10.

Scheiner R, Malun D (2001) Learning in honey bees with brain lesions: how partial mushroom-body ablations affect sucrose responsiveness and tactile antennal learning. *Animal Cognition* .

Scheiner R, Müller U, Heimburger S, Erber J (2003) Activity of protein kinase A and gustatory responsiveness in the honey bee ( Apis mellifera L.). *Journal of comparative physiology A* 189:427–34.

Scheiner R, Page RE, Erber J (2001a) Responsiveness to sucrose affects tactile and olfactory learning in preforaging honey bees of two genetic strains. *Behavioural brain research* 120:67–73.

Scheiner R, Page RE, Erber J (2001b) The effects of genotype, foraging role, and sucrose responsiveness on the tactile learning performance of honey bees (Apis mellifera L.). *Neurobiology of learning and memory* 76:138–50.

Scheiner R, Kuritz-Kaiser A, Menzel R, Erber J (2005) Sensory responsiveness and the effects of equal subjective rewards on tactile learning and memory of honeybees. *Learning & Memory* 12:626–635.

Scheiner R, Page RE, Erber J (2004) Sucrose responsiveness and behavioral plasticity in honey bees (Apis mellifera). *Apidologie* 35:133–142.

Schmuker M, Yamagata N, Nawrot MP, Menzel R (2011) Parallel representation of stimulus identity and intensity in a dual pathway model inspired by the olfactory system of the honeybee. *Frontiers in neuroengineering* 4:17.

Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annual review of psychology* 57:87–115.

Schwärzel M, Müller U (2006) Dynamic memory networks: dissecting molecular mechanisms underlying associative memory in the temporal domain. *Cellular and molecular life sciences* 63:989–98.

Smith BH, Abramson CI, Tobin TR (1991) Conditional withholding of proboscis extension in honeybees (Apis mellifera) during discriminative punishment. *Journal of comparative psychology* 105:345–56.

Smith B, Cobey S (1994) The olfactory memory of the honeybee Apis mellifera. II. Blocking between odorants in binary mixtures. *Journal of Experimental Biology* 108:91–108.

Smith BH, Huerta R, Bazhenov M, Sinakevitch I (2012) Distributed Plasticity for Olfactory Learning and Memory in the Honey Bee Brain In Galizia CG, Eisenhardt D, Giurfa M, editors, *Honeybee Neurobiology and Behavior*, chapter 6.1, pp. 393–408. Springer Netherlands, Dordrecht.

Smith BH, Menzel R (1989a) An analysis of variability in the feeding motor program of the honey bee: The role of learning in releasing a modal action pattern. *Ethology* 82:68 –81.

Smith BH, Menzel R (1989b) The use of electromygram recordings to quantify odourant discrimination in the honey bee, Apis mellifera. . *J.Insect Physiol.* pp. 369–375.

Smith D, Wessnitzer J, Webb B (2008) A model of associative learning in the mushroom body. *Biological cybernetics* 99:89–103.

Stollhoff N, Menzel R, Eisenhardt D (2005) Spontaneous recovery from extinction depends on the reconsolidation of the acquisition memory in an appetitive learning paradigm in the honeybee (Apis mellifera). *The Journal of Neuroscience* 25:4485–92.

Strube-Bloss MF, Nawrot MP, Menzel R (2011) Mushroom body output neurons encode odor-reward associations. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 31:3129–40.

Sutton RS, Barto AG (1990) Time-Derivative Models of Pavlovian Reinforcement In M G, J M, editors, *Learning and Computational Neuroscience: Foundations of Adaptive Networks*, pp. 487–537. MIT Press.

Szyszka P, Demmler C, Oemisch M, Sommer L, Biergans S, Birnbach B, Silbering AF, Galizia CG (2011) Mind the gap: olfactory trace conditioning in honeybees. *The Journal of Neuroscience* 31:7229–39.

Takeda K (1961) Classical conditioned response in the honey bee. *J Insect Physiol* 6:168 –179.

Tully T (1984) Drosophila learning: behavior and biochemistry. *Behavior genetics* 14:527–57.

Turner GC, Bazhenov M, Laurent G (2008) Olfactory representations by Drosophila mushroom body neurons. *Journal of neurophysiology* 99:734–46.

Wessnitzer J, Young JM, Armstrong JD, Webb B (2012) A model of non-elemental olfactory learning in Drosophila. *Journal of computational neuroscience* 32:197–212.

Wilson R, Turner G, Laurent G (2004) Transformation of olfactory representations in the Drosophila antennal lobe. *Science* 303:366–370.

Wright Ga, Choudhary AF, Bentley Ma (2009) Reward quality influences the development of learned olfactory biases in honeybees. *Proceedings. Biological sciences / The Royal Society* 276:2597–604.

Wright Ga, Smith BH (2004) Variation in complex olfactory stimuli and its influence on odour recognition. *Proceedings. Biological sciences / The Royal Society* 271:147–52.

Yarali A, Nehrkorn J, Tanimoto H, Herz AVM (2012) Event timing in associative learning: from biochemical reaction dynamics to behavioural observations. *PloS one* 7:e32885.

Young JM, Wessnitzer J, Armstrong JD, Webb B (2011) Elemental and non-elemental olfactory learning in Drosophila. *Neurobiology of learning and memory* 96:339–52.

Zhu C, Byrd R, Lu P (1997) Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM* pp. 1–17.