1.01 Introduction and Overview

R. Menzel, Freie Universität Berlin, Berlin, Germany

© 2008 Elsevier Ltd. All rights reserved.

1.01.1 Introduction

The central question of behavioral neuroscience is: What is the source of information that creates and controls perception, reaction, and action in animals and humans?

Since Plato and Aristotle, Western philosophy has found two opposing answers that still guide our thinking today. Idealism proposes preexisting information unfolding as the organism develops, whereas empiricism states that all information is gathered by the interaction of the organism with the environment. As behavioral science developed into its modern form in the second half of the nineteenth century, these two philosophical approaches provided the epistemological framework for two different experimental approaches to animal behavior. Ethology emphasized the preexisting information inherited from the evolution of the species, whereas behaviorism, including Pavlov’s physiology-driven approach, focused on collecting information through perception and action.

Reference to idealism and empiricism gave these two opposing approaches a strong theoretical backbone, turning anthropomorphic descriptions of behavior and pure collections of observations into hypothesis-driven science.

Ethology benefited from Darwin’s theory of evolution that provided a conceptual framework for the accumulation and transmission of information, while behaviorism gained from the strength of laboratory-based experimentation and suitability for formal descriptions (See Chapters 1.03, 1.06). Although the history of the two behavioral science disciplines is a success story, we also know of their limitations and their failure to capture the breadth of behavior.

The dualistic (and initially exclusive) conceptualization of the two forms of information that drive behavior is not adequate to explain brain functions. Regardless of the source of the information stored in the brain, it is expressed in properties of the brain, the wiring of neurons, and the communication between them. The two historical disciplines in behavioral science also failed to develop concepts that included brain functions. Ethology relied too heavily on simple-minded models of brain function, whereas behaviorism eliminated any reference to the brain.

In the end, behavior needs to be explained by underlying body functions, with the brain as the most important player in this game. Where are we now in such an attempt? This volume deals with behavior, theory, and system functions of the brain.

Although cognitive neuroscience and the “emergent new science of mind” (Kandel, 2006) offer novel levels of integration between behavioral science and neuroscience, we still have a long way to go. The chapters in this volume should contribute to this exciting endeavor.

1.01.2 Biology of Learning and Memory: The Value of a Comparative Approach

Learning from experience is a property embedded into the survival strategies of all animals living in natural surroundings. Animal species live in different ecological niches, are equipped with different sensory and motor capacities, and communicate differently with other individuals of the same species and with other species (See Chapters 1.17, 1.20, 1.21, 1.22, 1.26, 1.31). They also come with different nervous systems, which can be large or small, and some are highly centralized, while others have several rather separate ganglia.

There is no one model animal for this research.
endeavor, and different animal species have different advantages. One species can be reared more easily in the lab, another has a better-worked-out toolbox of molecular genetics, another provides for large and identifiable neurons or allows recordings from multiple neurons simultaneously over long periods of time, while another has already been analyzed in a wide range of behavioral tests. The main goal of this research is to unravel the general rules and species-specific adaptations in selecting relevant information, adding it to existing knowledge, storing it such that passing time does not eliminate it, and making it available for better-adapted behavioral acts in the future. Comparative studies provide the tool for identifying generalities and specificities.

Observing animals in their natural habitat can suggest relevant research questions (See Chapters 1.17, 1.20, 1.21, 1.22, 1.23, 1.25, 1.26, 1.29). While we have ideas about what is worth observing and measuring, we need to be open-minded about unexpected outcomes, as these are often the discoveries that propel research. Food-storing behavior in birds and mammals (See Chapters 1.22, 1.23), communication via gestures and/or sound (See Chapters 1.16, 1.17) or by ritualized movements in bees (See Chapters 1.12, 1.25, 1.29), learning during courtship in Drosophila (See Chapter 1.28), and navigation (See Chapters 1.20, 1.21, 1.25, 1.26) are examples, and many more will be described in this volume.

The animal species studied are varied: Worms (Caenorhabditis elegans; See Chapter 1.32), mollusks (Aplysia, Limax, Lymnaea, Hermissenda, See Chapter 1.30), cephalopods (octopus, Sepia, Loligo; See Chapter 1.31), insects (Drosophila, See Chapter 1.28; the honeybee Apis, See Chapters 1.12, 1.25, 1.29), rodents (e.g., mice, rats, squirrels, chipmunks, deer mice, Merriam’s kangaroo rats; See Chapter 1.22), birds (e.g., Clark’s nutcrackers, Eurasian nutcracker, marsh tits, black-capped chickadees, pinyon jays, Mexican jays, Western scrub jays, willow tits, crested tits, many song-learning bird species, the oscines of the Passeriform birds; See Chapters 1.17, 1.22, 1.23), and primates (chimpanzees, orangutans, macaques, vervet monkeys, baboons, capuchin monkeys, and pygmy marmosets; See Chapter 1.16), including human beings. Each of these species provides us with the opportunity to discover novel ways of solving similar problems brought about by an ever-changing environment and to unravel general strategies by comparison.

The transition from the natural habitat to the laboratory is an essential step in hypothesis-driven behavioral research, but is by no means a simple step (See Chapter 1.11). We cannot expect to get full control over the animal, which is sometimes a misleading assumption in some of the behavioral studies (See Chapters 1.03, 1.06, 1.10, 1.18). However, the history of an individual’s experience can be traced more accurately in the laboratory, and the proper control experiments can be established. The chapters in this volume provide ample evidence for successful transition from the natural environment to the laboratory.

Often an additional and more critical step is required. Animals need to be constrained for physiological measurements, or they are genetically manipulated to isolate cellular components of neural function. It is essential to remember that constrained animals or genetically manipulated animals are no longer the living creatures that we observed in their natural environment, and they are not even those seen in the laboratory behavioral tests. It might not matter so much whether a nutcracker caches a nut in the bark of a tree or in a plastic Lego building block, but it will make a great deal of difference whether a chimpanzee learns to move between tree branches or is sitting in a training chair connected to amplifiers while performing stereotypical arm movements. It is true that in most cases we do not yet have alternatives attempting to relate neural function with behavioral performance, but while presently there are no better experimental tools of neural recording, we must not forget the distance between the natural conditions and the experimental surroundings where we collect data.

Invertebrates such as the nematode C. elegans (See Chapter 1.32), the mollusks Limax or Aplysia (See Chapter 1.30), the fly Drosophila (See Chapter 1.28), and the bee Apis (See Chapters 1.25, 1.27, 1.29) are of particular value in laboratory settings since the transitions from natural to laboratory to experimentally interfering conditions presumably impact them less. Intermediate transition steps from natural to laboratory conditions can be made more easily, and the behavior of these invertebrates appears to be controlled more strongly by innate components. Nevertheless, transgenic nematodes and flies are not normal animals with just one isolated function that has been modified. It is, therefore, very advantageous that transgenes in C. elegans and Drosophila can be switched on and off rather quickly, and even more important, can be genetically rescued, which allows us to test the isolation of the targeted effect very carefully. Transgenic mice are by far more complicated to analyze, and special care must be taken in interpreting their behavioral alterations.
Many questions about learning and memory cannot be moved into the laboratory, and these may often be the particularly interesting questions (See Chapters 1.16, 1.22). This has two consequences: (1) The data are correlational in nature as control groups often cannot be studied or serve as partial controls and animal manipulations are very difficult or impossible, and (2) recording brain functions is difficult or impossible. These limitations should not reduce our efforts to collect data under natural conditions, as these data are essential for future laboratory studies and for comparative studies in humans.

A comparative approach should include human beings, and the motivation of many animal studies is to better understand humans. This is justified if appropriate caution is taken and the general limitations of a comparative approach are observed. Both ethology and behaviorism carry their historical burdens regarding inappropriate generalizations between animals and humans (See Chapters 1.03, 1.10), but cognitive neuroscience offers tools and strategies that help to guide such comparisons. If processes and mechanisms have been identified that apply across animal species, they are less likely to be species-specific adaptations and can safely be generalized to humans (e.g., cellular and molecular processes of neural plasticity as discussed elsewhere in this work, or basic rules of associative learning (See Chapters 1.05, 1.06, 1.09, 1.10, 1.11, 1.18)). The involvement of homolog brain structures for related forms of learning and/or memory formation are strong hints for homolog function. The hippocampus (e.g., in the case of spatial learning and episodic-like memory) and the amygdala (e.g., in fear learning) are two examples, and many more will be found in the chapters of this volume (See Chapters 1.14, 1.15, 1.21, 1.22, 1.23). Comparison between animals with very different brain structures (e.g., mammals and insects) is much more difficult, and often no more than analog functions can be assumed (e.g., navigation according to a geometric representation of space, different forms of memories according to their time course, and susceptibility to experimental interference).

One of the most important and controversial issues related to comparison between animals and humans relates to language and self-awareness (See Chapters 1.15, 1.16, 1.23, 1.37, 1.38). Although language acquisition has deep biological roots, the relation to nonverbal or acoustic communication in monkeys is not yet understood (See Chapter 1.16). The neural requirements of self-awareness exist in animals (See Chapters 1.37, 1.38), but it is not clear whether additional neural functions are required for the human form of self-awareness. The case of episodic memory, as discussed in several chapters (See Chapters 1.21, 1.22, 1.23), is a particularly interesting example because essential features of knowledge about what happened, when it happened, and where it happened exist at different degrees of complexity in many animal species (e.g., in the honeybee; See Chapters 1.12, 1.25, 1.29). Food-storing birds appear to relate these memories to themselves and appear to expect to find food at that location in the future (See Chapter 1.23), showing a capacity that is close to personal recollection in humans. Salwiczek et al. (See Chapter 1.23) call this memory episodic-like and see a gradual, rather than a principle, difference with the introspective experience of mental time travel in humans. This pragmatic approach might be exemplary in the sense that other human mental functions could also be broken down into additive features, which could then be tested for their existence in animals in various combinations and complexities. However, the demonstration of the existence of the components does not prove that the full function of a cognitive faculty as observed (or personally experienced) in humans exists in a particular animal species. Nevertheless, the strength of this approach lies in the assumption that there are no categorical differences between animals and humans, and gradual differences can be traced to different performances according to the complexity of the elements found. An example could be dance communication in honeybees (See Chapters 1.12, 1.25, 1.29). The bee communicates a location, and depending on the context, the dance might indicate a feeding place, a water or resin resource, or a new nest site. Although the communication process is symbolic and has a vocabulary (although a very reduced one) and a form of syntax (context-dependence), it does not qualify as a language because it lacks essential features, for example, semantics and grammar. One might call it language-like, as one might categorize other symbolic indexical forms of communications, but the point is that a research program can be set up by this decomposition strategy which allows scientists to search for the related neural processes of the components rather than the mental faculty as a whole. It appears to me that a similar research strategy is proposed by Changeux and Dehaene (See Chapter 1.38) in their attempt to decompose conscious processing and learning. They propose a unified or global workspace for the neural synthesis of past, present, and expected experiences. Such research approaches are promising because they
avoid the epistemological deadlock connected with the preoccupation of a categorical separation between animals and humans (MacPhail, 1998) or the assumption of equality between the animal mind and the human mind (Griffin, 1984).

1.01.3 Theories, Processes, and Mechanisms

Animal learning theory has been a rich research area over the last 60 years or so, and we may ask whether some of its concepts might join with physiological studies for a better understanding of the underlying processes. Theories derived from associative forms of learning have been elaborated the most (See Chapters 1.03, 1.06, 1.09, 1.10, 1.18), and it appears that three concepts are most useful in a search for functional implementations: associative strengths, associability, and prediction error (Dickinson, 2007).

1. Associative strength between two elements (stimulus or response) depends on the history of experience and the stimuli/responses involved and controls both acquisition and retrieval of memory. Although different behavioral theories compete for the best way of capturing the essence of associative strengths (e.g., Rescorla and Wagner, 1972; Pearce and Hall, 1980; Bouton, 1994; See Chapters 1.03, 1.06, 1.09), it is not yet clear whether unidirectional or bidirectional associations predict the relations between the elements. Neuroscientists are more than prepared to absorb this concept and translate it into processes of neural plasticity. Donald Hebb (1949) proposed such a neural implementation, and it is widely accepted that synaptic strength is closely related to associative strength (See Vol. 4). Long-term potentiation and long-term depression are processes that are based on the accurate timing of neural activity in the pre- and postsynaptic elements of neural nets (See Chapter 1.34).

The coincidence of spike activity as a means of modulating synaptic efficacy appears to play a role not only between pairs of pre- and postsynaptic neurons, but also in networks of many neurons. Singer (See Chapter 1.37) points out that coherence of spike activity is an essential feature of cortical nets in up- and downregulation of learning-related neural plasticity. It will be important to show that spike synchrony in biological networks is an emergent property similar to artificial networks (See Chapter 1.34) and to establish the causal relationship between these global network characteristics and learning. Since small networks composed of identified neurons do not depend on spike coherence in a global sense to establish associative changes in synaptic efficacy (e.g., in mollusks; See Chapter 1.30), it will be interesting to search for additional qualities of synchronizing neurons. Such additional qualities could lie in the fact that the three components of memory (formation, retrieval, and consolidation) are so tightly connected that only under conditions of synchronized activity are all three memory components activated. New contents can only be stored in distributed brain regions which jointly reorganize the network according to the new information.

2. Associability is another concept developed in behavioral learning theory that promises to be useful in neural studies. The concept captures the properties of the stimuli and/or outcomes that determine associative strengths as they are reflected in the salience of the stimulus, the predictability or surprise value of a stimulus, or the outcome. Cognitive dimensions of operant learning (See Chapters 1.06, 1.10) or perceptual learning (See Chapter 1.07) involve attention as a critical parameter of learning (See Chapter 1.13), a parameter that can be traced to particular structures (e.g., cholinergic projections from basal ganglia, amygdala, and the septohippocampal system).

3. Prediction error. Learning theories state that learning occurs as long as the outcome of a behavior is not fully predicted, and thus the deviation of the expected from the experienced outcome changes the current associative strengths. Behavioral theories differ with respect to their assumption of whether the error affects associative strength directly (e.g., Rescorla and Wagner, 1972; Mackintosh, 1975) or indirectly (Pearce and Hall, 1980). The implementation of the prediction error into machine learning (Sutton and Barto, 1990) has been very successful (See Chapters 1.34, 1.36), and strong neural correlates exist: for example, the neural properties of reward neurons (dopamine neurons of the mammalian ventral tegmentum (Schultz, 2006) and octopamine VUMmx1 neurons in the honeybee brain (Hammer, 1992; Menzel and Giurfa, 2001)).

When dealing with learning, several chapters in this volume draw their underlying concepts from both behavioral and neural data to document that the strongest expectations for an understanding of processes and mechanisms come from collecting experimental data from both behavioral and neural studies (Dickinson, 2007). How far will such a hybrid theory and understanding lead us?
 Forty years ago, Kandel and Spencer wrote a seminal paper entitled “Cellular neurophysiological approaches in the study of learning,” calling for a novel approach in translating basic psychological concepts of learning into strategies for the search for their neural implementations (Kandel and Spencer, 1968). Less than 20 years later, Hawkins and Kandel (1984) presented a first review on their finding on *Aplysia* associative and nonassociative learning and derived neural components comprising a cellular alphabet of learning (See Chapters 4.01, 4.02). This strategy has turned out to be most successful in localizing in space and time neural events induced by learning. It appears that the associative events are distributed, multifaceted, and dependent both on innate predispositions and earlier learning. *Drosophila* provides a particularly carefully studied case (See Chapter 4.07). Different neural structures are involved in learning the same odor by reward or punishment, and short- and long-term memories of the same content reside in different neural nets. Localizing the memory trace is an important step in a functional analysis.

A major unresolved issue in both behavioral and neural studies is the relationship between learning with and without external reinforcing or evaluating stimulus. As pointed out above, concise behavioral theories have been developed for Pavlovian and instrumental conditioning, but perceptual learning (See Chapter 1.07), navigational learning (See Chapters 1.12, 1.20, 1.25), and interval learning (See Chapter 1.19) provide cases in which no obvious external reinforcer may be present. Is associative learning a special case of a more general form of learning (the learning of temporal sequences; See Chapter 1.12), or is every kind of learning associative? Does an internal reinforcer provide the evaluating function in the latter forms of learning? Learning theory has not settled the debate, and it might well be that functional analysis will show that internal reinforcing circuits are active at the proper time when animals learn by observation. An important component in such forms of learning is attention (See Chapters 1.13, 1.36), as is assumed in a modeling study of navigational learning in the honeybee (Montague et al., 1995).

Only selectively attended stimuli are learned. Most importantly, modulatory circuits that appear to be involved in coding evaluating stimuli also participate in selective attention. It will be necessary to build conceptual bridges between the concept of associability as developed in theories of associative learning and the evaluating property of directed attention as described in observational learning. Further advances will only be made with the combination of behavioral and neural approaches.

### 1.01.4 What Is Memory and What Is a Memory Trace?

The many facets of memory are reflected in the many terms used to capture them (See Chapters 1.02, 1.04). Are there 256 different kinds of memory, as Tulving (1972) asked? Irrespective of whether we divide up memories according to time, cellular mechanisms, brain structures involved, categories of contents, type of learning, or type of retrieval, we always imply that memory directs behavior via the process of retrieving information. As pointed out at the beginning of this chapter, brains are equipped with information before, and independent of, acquired information. Thus the content of memory provides a knowledge base for behavioral guidance (including perception, planning, expecting, and thinking), and splitting it up may obscure the basic and unifying property of memory. One question that needs to be asked, then, is: How do we go about measuring the knowledge stored in memory? We do not know, and this ignorance might be one of the reasons why so much emphasis is placed on the need to define memory by retrieval processes (See Chapters 1.02, 1.04, 1.12, 1.14, 1.15, 1.24). As long as measurement of memory content is based only on retrieving it from memory, we will not be able to separate stored memory from used memory.

Nadel (See Chapter 1.04) quotes from Aristotle’s *Ars Memorativa*: “It has already been stated that those who have a good memory are not identical to those who are quick at recollecting. But the act of recollecting differs from that of remembering, not only chronologically, but also in this – that many of the other animals (as well as man) also have memory, but of all that we are acquainted with, none, we venture to say, except man, shares in the faculty of recollection.” Indeed, the distinction between memory and recollection is multifaceted (See Chapters 1.02, 1.04, 1.05, 1.14, 1.15, 1.24), and one of the most important distinctions relates to memory formation versus memory retrieval. Behavioral measures, as well as human subjective introspection, reach memory via retrieval, but the postlearning reactivation process can work only if learning left traces in the form of an engram. Since the process of memory formation is
not directly accessible to behavioral studies, it has been seriously questioned from a behavioral analytical perspective, in terms of whether it makes sense to distinguish between memory as an entity independent of retrieval (See Chapters 1.04, 1.05). The notion of a physical memory trace, independent of its use, however, is a central presumption in neuroscience. Indeed, only when neurologically related interference procedures were introduced into memory research did a clear separation between memory formation and memory retrieval become possible. The key discovery in this context was the consolidation process.

Does a memory exist if it is not retrieved? This question is addressed in several chapters in this volume, and rather diverse opinions are expressed. If the knowledge stored in memory does not guide behavior, a behavioral biologist cannot know whether memory exists (and may thus define memory by its retrievability). But a neuroscientist cannot help but assume that the knowledge stored in memory continues to exist during time periods when it is not retrieved, because the physiological measures of memory are independent of whether the animal performs the corresponding behavior. The concept of memory consolidation is essential in this debate. Hermann Ebbinghaus (1964) described a fast and a slow component in forgetting, and William James (1890) proposed that these may be related to two forms of sequential memories: Primary and secondary memory. The concept of consolidation as a time-dependent process following learning was introduced by Müller and Pilzecker (1900) on the basis of their finding that new learning interfered with the formation of recently acquired memory for short, but not for long intervals. At this stage of analysis, a separation between an internal, time-dependent, and self-organizing process of memory formation and retrieval of memory was not possible, but when experimental interference was introduced and neurological cases of retrograde amnesia were analyzed, strong arguments in favor of an independent engram-building process could be presented (See Chapters 1.14, 1.15). However, the situation is not as simple as was believed (See Chapters 1.04, 1.05). For example, amnesia-inducing procedures could have led to competing learning processes. Irrespective of the unresolved questions in separating memory formation and memory retrieval processes, the body of evidence is overwhelming, proving that neural traces are indeed induced by the learning process independent of retrieval, and consolidation has a physical basis in the structuring and restructuring processes of neural net properties.

Procedures interfering with ordered neural activity or cellular metabolism during periods of consolidation induce retrograde amnesia. Memory gets better over time, even when it is not used. Sleep phases strengthen the consolidation process (Born et al., 2006) and are related to repetition of content-specific patterns of neural activity (Wilson and McNaughton, 1994). It appears to me that the debate about the nature of the memory trace (See Chapter 1.04) will continue as long as we cannot read the encoding processes and directly measure knowledge stored in neural nets. Once we can show these in suitable animals such as Drosophila, we will probably discover that, in addition to the constructive processes of reactivating memory and using its content, there is an essential component that exists independent of the reactivation process. Whether we like to call this lasting component memory is a question of definition.

Reactivation of memory leads to new learning and its subsequent consolidation processes (See Chapter 1.09). Only recently has neuroscience become interested in the mechanistic aspects of extinction learning and memory formation. The phenomena subsumed under the term reconsolidation provide case studies (See Chapters 1.24, 1.27). Reconsolidation refers to the effect that retrieving memory may lead to cue-dependent amnesia if the retrieval process is followed by treatment with an amnestic agent. What are these learning and reconsolidation processes? Does reactivation indeed make the old memory trace vulnerable to amnestic interference, indicating that new learning overwrites old memory, or do the learning processes involved in memory reactivation induce parallel consolidation processes that reflect the addition of a new memory trace to the existing one? The ongoing debate reflects the same dilemma addressed above. Our inability to measure knowledge as stored information directly restricts our mechanistic analysis to global and indirect arguments. Once again, behavioral analysis needs to be combined with fine-grained neural analysis addressing the critical question much more directly at the level of the neural elements of the engram.

What might be a suitable strategy toward a direct reading of knowledge? A first step should be to develop criteria that allow us to identify and localize a memory trace. Heisenberg and Gerber (See Chapter 1.28) address this question by defining four essential requirements of a memory trace:
1. Neuronal plasticity occurs in particular neurons that are localized and identified, and these neurons are essential for a particular kind of memory.

2. The neuronal plasticity in these neurons is necessary for this particular memory context.

3. Memory cannot be expressed if these neurons cannot contribute during retrieval.

4. Memory cannot be established if these neurons do not receive the required input for the memory content to be stored.

Note that this checklist of experimental procedures does not yet provide us with access to information stored in the memory trace, but we can hope that in a next step a localized and thus-characterized memory trace will be accessible to the really important question of how neural circuits encode and store particular pieces of information. So far only one organism, Drosophila, offers the opportunity to localize and characterize a memory trace at the level of cellular resolution, and, indeed, in applying this strategy to rather small neural circuits it was found that traces for short-term memory of an olfactory discrimination task and long-term forms of the same memory content appear to be localized in different, probably partially overlapping neural circuits. Furthermore, it was found that memory traces of appetitively or aver-sively evaluated stimuli of the same kind occupy different but partially overlapping neural circuits (See Chapter 1.28).

A whole battery of highly sophisticated molecular–genetic tools are available to measure the spatial–temporal patterns of memory traces in selected neurons and neural nets of the Drosophila brain (Keene and Waddell, 2007). Reading the dynamics of the neural elements during the learning process (i.e., consolidation and retrieval under conditions in which the animal tells us via its behavior whether it perceives, attends, and retrieves) will help us understand at least part of the knowledge stored in memory. How close are we then to direct knowledge reading?

Localizing and characterizing the memory trace by applying correlation analysis is the mainstream of the neuroscience approach today (See the volumes edited by Sweatt and Eichenbaum). Correlating elemental with system properties is an important step in any mechanistic analysis. The next step will be to establish closer, possibly causal, links following the strategy outlined by Heisenberg and Gerber (See Chapter 1.28). The tools also exist for the worm C. elegans (See Chapter 1.32) and are becoming available step by step for other species (e.g., the mouse). The hunt for direct knowledge reading will be embedded in a concerted approach to understanding the workings of neural nets and the brain as a whole.

### 1.01.5 The Engineer’s Approach to Learning and Memory

Engineers compose and biologists de-compose, so a combination of these two strategies should be favorable to the study of a complex system such as the brain. Constructive thinking in theoretical neuroscientists is inspired by rules derived from behavioral studies (e.g., Hebb’s rule), by the morphology of brains and the connectivity patterns of neurons (e.g., the matrix-like connectivity in the hippocampus), by the functional properties of neurons (e.g., synaptic plasticity), and by theoretical concepts developed independently from, but motivated by, thoughts about how the brain might work (e.g., autoassociative or attractor networks). Our volume contains chapters dealing with all of these aspects (See Chapters 1.33, 1.34, 1.35, 1.37, 1.38). Irrespective of the intellectual pleasure one experiences when thinking about theoretical neural nets, one might ask how the joint efforts propel our understanding. I see the following points that are also well-illustrated in the respective chapters:

1. Hypothesis-driven research like ours requires well-formulated concepts and hypotheses. Theories developed for neural nets shape these concepts and allow us to formulate predictions (See Chapters 1.33, 1.35).

2. The analysis of the vast amounts of data collected by anatomical, electrophysiological, optophysiological, and molecular studies requires the contribution of theoretical neuroscientists to extract relevant information and interpret it (See Chapters 1.34, 1.37).

3. There exists no concise theory of the brain. Global brain functions need to be constructed from elemental and network functions and implemented into a model (e.g., the neuronal workspace model of Changeux and Dehaene, See Chapter 1.38).

At any of these levels of a modeling approach, one has to decide what is considered an essential feature and which of the many characteristics of the neurons, their connectivity at the local and the global level, are implemented or not. Should one use simplified integrate-and-fire neurons or Hodgkin-Huxley-type neurons? Should the model care about the real gestalt of neurons or not? How seriously should one take the
neuroanatomical data on local and global connectivity? These and many other decisions are hard to make, and different choices produce serious debates about the suitability of these models. There are many measures of suitability: Are experiments stimulated, predictions offered, and interpretations of data supported or rejected? Five chapters (See Chapters 1.33, 1.34, 1.35, 1.37, 1.38) provide a range of examples where strong arguments can be presented for the suitability of the respective models and experimental approaches are suggested. Other examples are given in Vol. 4 that deal with small neural nets partially composed of identified neurons (See Chapter 1.30). Indeed, small biological neural nets (e.g., the stomatogastric ganglion in the lobster) have been successfully modeled, and the models in the electronic version were directly hooked up with the biological neural net to analyze the contribution of certain cellular properties (Golowasch et al., 1999). These approaches have been applied in the search for the neural implementation of operant learning in the buccal ganglion of Aplysia (Brembs et al., 2002; Vol. 4; See Chapter 4.10).

Given the technological advances with the expression of light-driven conductances in specified neurons (e.g., channel rhodopsin), similar analyses will be possible (e.g., in Drosophila, See Chapter 1.28), which emphasizes the need for theoretical concepts and models of the respective networks. Ultimately, models of neural function should also predict behavioral outcomes. Singer (See Chapter 1.37) makes the case for the role of spike synchrony and oscillatory spike patterns in memory formation. Koene and Hasselmo (See Chapter 1.35) formulate predictions for the role of theta rhythm in the hippocampus for memory formation and retrieval, and Rolls (See Chapter 1.33) explicitly characterizes the properties of the autoassociative network of the hippocampus for behavioral phenomena such as completion and graceful degradation. It is to be expected that the success of the combined theoretical and experimental approach will make modeling an indispensable part of the search for the memory trace.

1.01.6 Conclusion

Curiosity-driven behavioral studies, theory-guided laboratory behavioral experiments, and modeling of neural functions define a unique workspace in the search for the engram. Joining forces will help, and the chapters presented here will hopefully facilitate communication between these disciplines. The task is indeed demanding, because the goal will not only be to localize and characterize the memory trace, but to measure the knowledge stored in the memory trace independent of and in addition to the behavioral read-out process.

References

Müller GE and Pilzecker A (1900) Experimentelle Beiträge zur Lehre vom Gedächtnis. Z. Psychol. 1: 1–288.
Rescorla RA and Wagner AR (1972) A theory of classical conditioning: Variations in the effectiveness of reinforcement