

Toward a Science of Computational Ethology

David J. Anderson^{1,2,4,*} and Pietro Perona^{3,4,*}

¹Division of Biology and Biological Engineering

²Howard Hughes Medical Institute

California Institute of Technology, Pasadena, CA 91125, USA

³Division of Engineering, California Institute of Technology, Pasadena, CA 91125, USA

⁴Co-senior author

*Correspondence: wuwei@caltech.edu (D.J.A.), perona@caltech.edu (P.P.)

<http://dx.doi.org/10.1016/j.neuron.2014.09.005>

The new field of “Computational Ethology” is made possible by advances in technology, mathematics, and engineering that allow scientists to automate the measurement and the analysis of animal behavior. We explore the opportunities and long-term directions of research in this area.

Nothing enters our minds or determines our actions which is not directly or indirectly a response to stimuli beating upon our sense organs from without. Owing to the similarity of our construction and the sameness of our environment, we respond in like manner to similar stimuli, and from the concordance of our reactions, understanding is born.—Nikola Tesla (“A Machine to End War”)

Contempt for simple observation is a lethal trait in any science.—Niko Tinbergen (“On Aims and Methods of Ethology”)

Introduction

Understanding how the brain works is one of the most fascinating and challenging problems facing 21st century science. Brain function involves sensory perception, emotion, cognition, learning and memory, and motor control. All of these functions serve ultimately to control an animal’s behavior. Behavior allows individuals and groups of animals to adapt to a changing environment on a timescale much faster than evolution by natural selection. It is that adaptive function, the “selection for flexibility,” that has driven the rapid evolution of brains across phylogeny. Accordingly, if we want to understand how the brain works, it is essential to consider its functions in the context of behavior (Dickinson et al., 2000; Frye and Dickinson, 2004). That means we need ways of quantitatively and accurately measuring behavior, in all its richness and complexity.

The biological study of animal behavior, including its phenomenological, causal, ontogenetic, and evolutionary aspects, is a discipline known as ethology (Tinbergen, 1951, 1963). Ethology has a rich tradition, going back to the work of Konrad Lorenz, Karl von Frisch, and Niko Tinbergen, who received the 1973 Nobel Prize for their fundamental contributions. While ethology is a multifaceted discipline, at its core is the description and characterization of behavior, typically of intact freely moving animals in their natural environment. At its inception, such descriptions were primarily qualitative in nature. Over the last 40–50 years, however, the analysis of behavior has become increasingly quantitative. For example, while watching an animal behave, a human observer, rather than simply writing down a description of what he or she sees, will score incidents of specific behaviors defined according to certain generally accepted criteria. This

produces a numerical, rather than a purely written, description of the behavior or behaviors being observed.

While such quantification was at first typically performed using the traditional pencil, paper and stopwatch, advances in technology have facilitated this laborious process. There are now computer-assisted video analysis tools that allow user-defined behaviors to be scored with a keystroke, frame-by-frame. From such measurements, an observer can compute various statistics, such as the frequency of a given behavior, the latency to initiate a behavior, the duration of a behavioral episode, and the relative proportion of different behaviors or activities. Such measurements can be compared with simultaneous electrophysiological recordings of neuronal activity, permitting quantitative correlations between behavior and neuronal spiking activity.

Behavior has also been quantified using instrumental methods such as audio recording (Insel and Winslow, 1991; Williams, 2004). In the field of birdsong research, for example, sonograms derived from such recordings have been a key quantitative readout for experimental manipulations of the underlying brain circuitry (Moorman et al., 2011). In the field of chronobiology, measurements of locomotor activity in *Drosophila* based on IR beam breaks (Hamblen et al., 1986; Tataroglu and Emery, 2014) have been essential to the identification of genes and circuits that control circadian rhythms (Nitabach and Taghert, 2008). Quantification of turning behavior during flight using torque sensors (Reichardt and Poggio, 1976; Heisenberg and Wolf, 1984) or wing-beat detectors (Götz, 1987) has been valuable for studying orientation control and optomotor responses in *Drosophila*. The ability to correlate neuronal activity with quantitative analysis of behavior has been essential to the development of the field of neuroethology and has led to the development of important concepts such as the “command neuron” (Bentley and Konishi, 1978).

Neuroscience is in the midst of a revolution fueled by spectacular new technologies for mapping, monitoring, and manipulating neural activity based on genetic targeting of specific neuron subtypes (Luo et al., 2008) (Figure 1 and Table 1). These methods afford the ability to move beyond correlation to establishing causal relationships between neural circuit activity and behavior. New tools, such as optogenetics (Yizhar et al., 2011) and pharmacogenetics (Farrell and Roth, 2013),

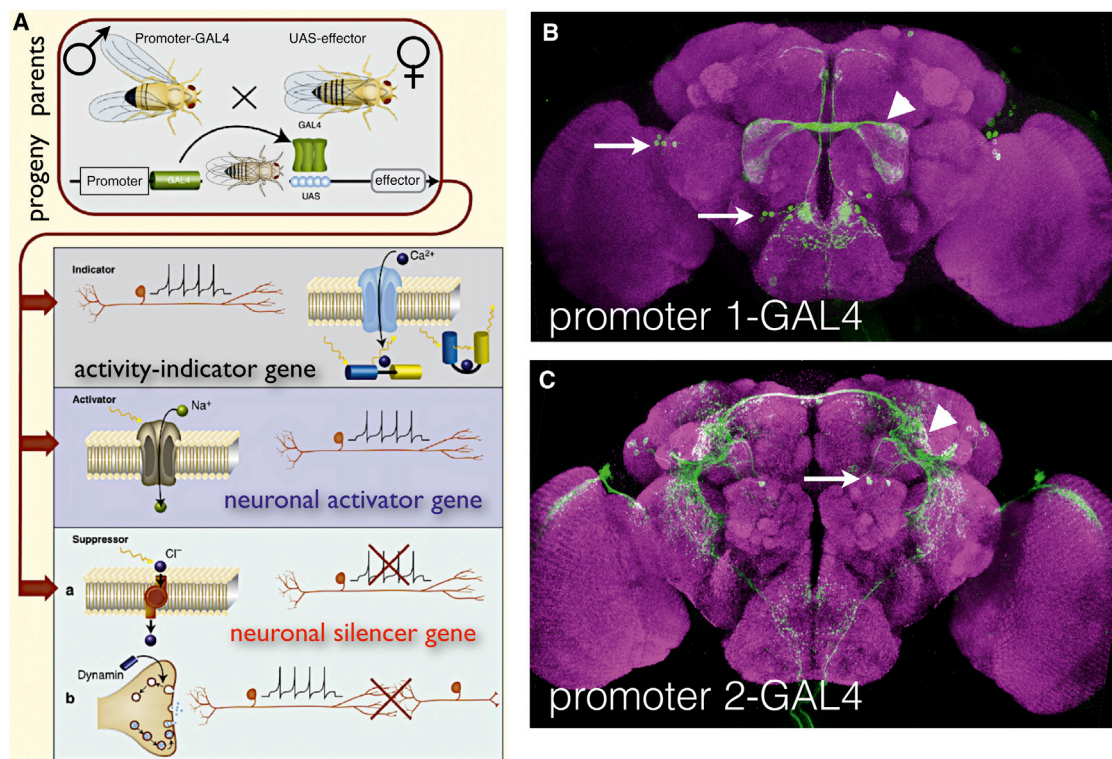


Figure 1. Mapping, Monitoring, and Manipulating Specific Neuron Populations in *Drosophila* Using the GAL4/UAS Binary System

(A) Upper: parental fly lines that contain a promoter-GAL4 or UAS-effector transgene (Table 1) are crossed to yield progeny that express GAL4, and therefore the UAS-effector, in a subpopulation of neurons defined by the specificity of the promoter (B and C). Lower: the effector can be an indicator to monitor the activity of the neurons, or ion channels or other genes to activate or silence the neurons. Reproduced with modifications by permission from Borst (2009). (B and C) A fluorescent marker allows visualization of neurons identified by different promoter-GAL4 lines. Arrows indicate cell bodies (green dots); arrowheads indicate nerve fibers.

connectomics (Reid, 2012), and optical imaging of neuronal activity (Kerr and Denk, 2008), are transforming our ability to understand how neural circuits control sensory perception, cognitive processes, internal brain states, and behavior (<http://www.nih.gov/science/brain/2025/index.htm>).

Exploiting this transformative technology is, however, critically dependent on the ability to assess quantitatively, and with a high degree of spatiotemporal precision, the behavioral consequences of neural circuit manipulations. However, the technology for measuring behavior has not kept pace with the rapid development of these new methods; manual scoring of behavior is (with notable exceptions described below) still the dominant approach in the field. This has hampered progress in both understanding the neural circuit control of ethologically relevant behaviors and in using behavior as a “read-out” for manipulations aimed at uncovering fundamental principles of neural circuit function. Reliance on human observation to score behavior imposes a number of limitations on data acquisition and analysis. These include:

It is slow. Human observer-based measurements of behavior, even when computer assisted, are extremely slow and labor intensive. (Annotating a video recording of an animal’s behavior typically takes a human observer 3× the length of the video.) This limits throughput and therefore reduces the number of different

experiments that can be performed, as well as sample size, thereby limiting statistical power and the reliability of results (Button et al., 2013).

It is imprecise and subjective. The scoring of behaviors by human observers is subjective, difficult to standardize, and therefore often inconsistent between different observers (Levitin et al., 2009). In our experience, two human observers will only agree 70% of the time in their analysis of the same data set. This makes it difficult to achieve consistency and reproducibility in behavioral analysis, both within and between laboratories.

It is low-dimensional. Even a single behavior, as will be discussed below, can be deconstructed into progressively finer components, whose granularity is ultimately limited by the structure and function of the animal’s motor system (Anonymous, 2007). While techniques exist for fine-scale analysis of motor behavior (Bui et al., 2013), measuring even a single behavior at multiple spatial and temporal scales simultaneously is currently challenging. Furthermore, due to the laborious nature of manual scoring, the number of different behaviors that can be measured concurrently is relatively small.

It is limited by the properties of the human visual system. Some aspects of this ability, such as distinguishing two animals even when they occlude each other, or recognizing a behavioral pattern that involves rapid and variable but nevertheless typical

Table 1. Genetic Manipulation of Neural Circuit Activity

Promoter	A segment of non-protein-coding DNA that contains information to cause (“drive”) expression of a gene in a particular cell or cell type. “Promoter” is often misused as shorthand for an “enhancer,” which is technically the more precise and correct term.
Effector	A gene that encodes a protein whose activity influences some aspect of cellular function. In the case of a neuron, this could be an ion channel, neurotransmitter transporter, receptor, etc. It can also encode a fluorescent protein that serves to reveal a cell’s anatomy (Figures 1B and 1C). Effectors can come from diverse species, e.g., jellyfish, algae, or bacteria.
Binary system	A general means of using a given promoter to express a given effector in a specific cell type, in order to functionally manipulate (e.g., excite or silence) that cell. The approach is “binary” in the sense that it typically comprises two inhibit distinct components, which must be genetically combined to achieve the desired manipulation. Binary systems are flexible, because they permit different combinations of promoters and effectors, either to perform the same functional manipulation (e.g., neuronal silencing) in two different cell types or different functional manipulations (e.g., silencing and activation) in the same cell type. The “GAL4/UAS” system is a binary system often used in <i>Drosophila</i> to manipulate different classes of neurons (Figure 1A).
Conditionality	A property that permits an effector to be turned on or off in a controlled manner, by manipulating parameters that influence protein function. Examples of such variables include temperature, light, drugs, or mechanical pressure. Conditional manipulations can be used to control the time at which a particular effector is turned on or off in a given cell type. Those parameters must have no independent influence on the behavior or interest. Not all effectors are conditional.
Thermogenetic	Temperature-dependent conditional control, e.g., of an ion channel or enzyme.
Optogenetic	Light-dependent conditional control, e.g., of an ion channel or enzyme.
Pharmacogenetic	Drug-dependent conditional effector control. Sometimes called “chemogenetic.”

actions (e.g., aggression), have not yet been surpassed by computers. However, computers may be able to spot new patterns that a human observer may have missed, due to inattention (see above), ascertainment bias, or timescales that exceed working memory. Simply put, humans are best at identifying and measuring behaviors that they expect to see. They are less good at discovering new ones.

It is limited by human language. The identification or classification of actions and activities by human observers cannot always be captured by formal verbal definitions. Some features of the operational definition of a behavior may therefore be difficult to communicate, even if they seem intuitively obvious to a given observer. This makes it difficult to train new observers to score behaviors in a manner that replicates the original criteria developed by the trainer.

It is dull. Spending hours and hours each day sitting in front of a video monitor is mind-numbingly boring, so an observer’s attention can easily drift. This not only increases the likelihood of error, but also the chance that interesting new behaviors will be missed. It also diverts effort from more creative uses of the human brain.

We argue here that a new interdisciplinary field, made possible by advances in mathematics, engineering, and computer science, is in a position to overcome these difficulties. This emerging field, which involves collaborations between physical scientists from these disciplines and biologists, has the potential to revolutionize the way in which we measure and model behavior. We call this field “computational ethology” (CE), to emphasize both its roots in the study of natural behavior in freely moving animals and the application of modern quantitative tools for measuring, describing, and analyzing behavior. It exploits recent advances in machine learning and machine vision (Table 2) to automatically identify and quantify instances of known, observer-defined behaviors, as well as to discover potentially new behaviors. One might wonder about the need of prepending “computational” to any field of science—it is obvious that computational tools are useful and will be gradually introduced

to almost any discipline. However, we feel that in the case of ethology, computational tools have the potential of transforming the study of behavior. With the new technology available for neural circuit analysis, this field should improve the ability to move beyond correlations to establish causal relationships between molecular and cellular mechanisms, circuit-level computations, and behavior. Here we describe the promise of CE and the challenges it faces to grow into a mature discipline.

The Impact of Automated Measurement: Lessons from Fly Social Behavior

Automating the measurement of behavior offers, in principle, solutions to many of the problems outlined above. Before delving into the details of how such measurements are achieved, however, it is useful to see how this approach has already impacted a particular field. The study of social behavior in *Drosophila melanogaster*, an area in which the authors have worked (Dankert et al., 2009), provides such an illustrative example. Social behaviors in the vinegar fly include both dyadic (pairwise) interactions such as male-female courtship and male-male aggression, as well as higher-dimensional interactions within large (>10) groups of flies. Innate social behaviors such as mating and fighting have, in particular, attracted attention because they offer the possibility of understanding how such behaviors are encoded in DNA (Dickson, 2008).

Increasing the dimensionality of behavioral analysis. Courtship in *Drosophila* is a complex process involving a series of discrete, stereotypic actions that progress in an ordered, but temporally variable, sequence (reviewed in Dickson, 2008; Yamamoto and Koganezawa, 2013) (Figure 2). Because of the labor involved in manually scoring each of these actions individually, courtship behavior has typically been quantified by means of a “courtship index” (CI) (Hall, 1994), which reflects the amount of time the fly spends engaged in any courtship-related behavior. Such a metric, however, conflates multiple courtship-related actions, such as chasing, licking, singing, or copulation, and gives them equal weight. The use of such a combined index can, therefore,

Table 2. Machine Vision and Machine Learning

Machine vision	Discipline concerned with enabling machines to “see,” similarly to biological organisms that use their eyes to measure properties of their environment. Machine vision systems consist of one or more cameras that are connected to a computer. Images collected by the camera(s) are transferred to the computer, where appropriate calculations are performed to extract the relevant information, e.g., the position, shape, and identity of objects in the scene. This information is then made available to the user or employed to control equipment such as an autonomous vehicle. A challenge in machine vision is computing invariant descriptors from images, i.e., descriptors of image content (e.g., the color of a surface, the shape of an object’s boundary) that are invariant with respect to irrelevant variations in the environment, such as the distribution and intensity of lighting.
Machine Learning	Discipline concerned with enabling machines to “learn,” instead of being programmed explicitly, to perform appropriate functions. Learning takes place with training from data and from human experts. In the case of computational ethology, a computer may be tasked with analyzing video and detecting specific animal actions, as specified by a human expert. For simple actions (e.g., “walking”), the behavior of the machine may be programmed explicitly by the expert in the form of rules (e.g., “if the fly moves faster than 0.1 mm/s, then it is walking”); in this case no machine learning is needed. For more complex actions (e.g., lunging) it is difficult for a human to program explicit rules to detect the action reliably. However, the expert will be able to provide examples of the action to be detected, in the form of short snippets of video where the action takes place (positive training examples) and short snippets where it does not (negative training examples). Machine learning algorithms will make use of the training examples and produce an appropriate “action classifier.” The performance of the action classifier may be assessed by testing it on new examples that were not used for training. Machine vision and machine learning researchers collaborate in designing vision systems that can learn from example.
Supervised learning	A classifier is trained to recognize specific patterns that are defined by a human expert. The expert provides the learning algorithm with a set of labeled patterns. The classifier is iteratively trained and tested, and more training examples may be added, until it achieves satisfactory detection of all true positive events (“recall”) and minimal contamination with false positive events (“precision”).
Unsupervised learning	A classifier is trained to detect and discriminate different patterns from unlabeled data, i.e., without information regarding how many and which patterns to look for. Unsupervised learning may lead to the discovery of previously unknown actions. Unsupervised learning is possible when the data cluster naturally, e.g., some actions may involve high-velocity motion and some may be carried out at low velocity.

confound the comparison of results between experiments, since similar CI values may reflect different underlying behaviors: a male fly that spends most of his time chasing a female could receive the same CI as a fly that spent the same amount of time singing. Recent studies have developed technology to automatically measure individual courtship-related behaviors (Dankert et al., 2009; von Philipsborn et al., 2011; Tsai and Huang, 2012; Asahina et al., 2014), using machine vision and machine learning-based approaches (see Figure 3 and Table 3). In time, one would ideally like to measure automatically multiple elements of the courtship behavioral repertoire with equal precision and accuracy, in order to be able to quantify the distribution, proportion, and relationship of each of these different actions over the course of an experiment, a sort of time-evolving ethogram.

Increasing the throughput of behavioral analysis. Automated analysis increases not only the dimensionality of behavioral measurements, but also its accuracy, consistency, objectivity, and—perhaps most importantly—experimental throughput. Until recently, *Drosophila* social behavior was (and largely continues to be) manually scored. This has limited analysis to dyadic interactions, simply because keeping track of 10–20 flies in an arena is virtually impossible for a human observer. However, recent studies have developed video tracking systems to measure the individual trajectories of dozens of flies simultaneously in an arena (Branson et al., 2009) (Figure 4). Analysis of these data has led to the discovery of new behaviors: for example, it was found that female flies in such groups occur within < 2 body

lengths of each other much less frequently than expected by chance, suggesting that they actively avoid proximity (Branson et al., 2009; Zabala et al., 2012). More recent efforts have allowed multiorganism tracking while identifying and discriminating individuals (Pérez-Escudero et al., 2014).

Even in the case of dyadic interactions such as courtship and aggression, the laboriousness of manual scoring has been a limiting factor in several respects. First, it makes it very difficult to perform high-throughput, unbiased “forward” genetic screens, for which *Drosophila* is an ideal system. Such screens are the best way that fundamentally new and unexpected discoveries can be made about genes or neurons that regulate behavior. Audio recordings have been used to conduct a systematic screen for neurons that control male courtship song (von Philipsborn et al., 2011), illustrating the power of such an approach. But automated measurements of additional aspects of courtship behavior (von Philipsborn et al., 2011; Tsai and Huang, 2012; Asahina et al., 2014), as well as of other social behaviors, such as aggression (Dankert et al., 2009), could greatly increase the applicability of this approach.

Automated analysis also affords greater depth, scope, rigor of experimental design, and statistical power in “low-throughput” hypothesis-driven research. It allows more hypotheses to be tested, more variables to be explored, and more controls to be included and enables larger sample sizes to achieve greater statistical power (Button et al., 2013). A recent study of aggression in *Drosophila* (Asahina et al., 2014), which utilized automated behavioral analysis (Dankert et al., 2009), illustrates the

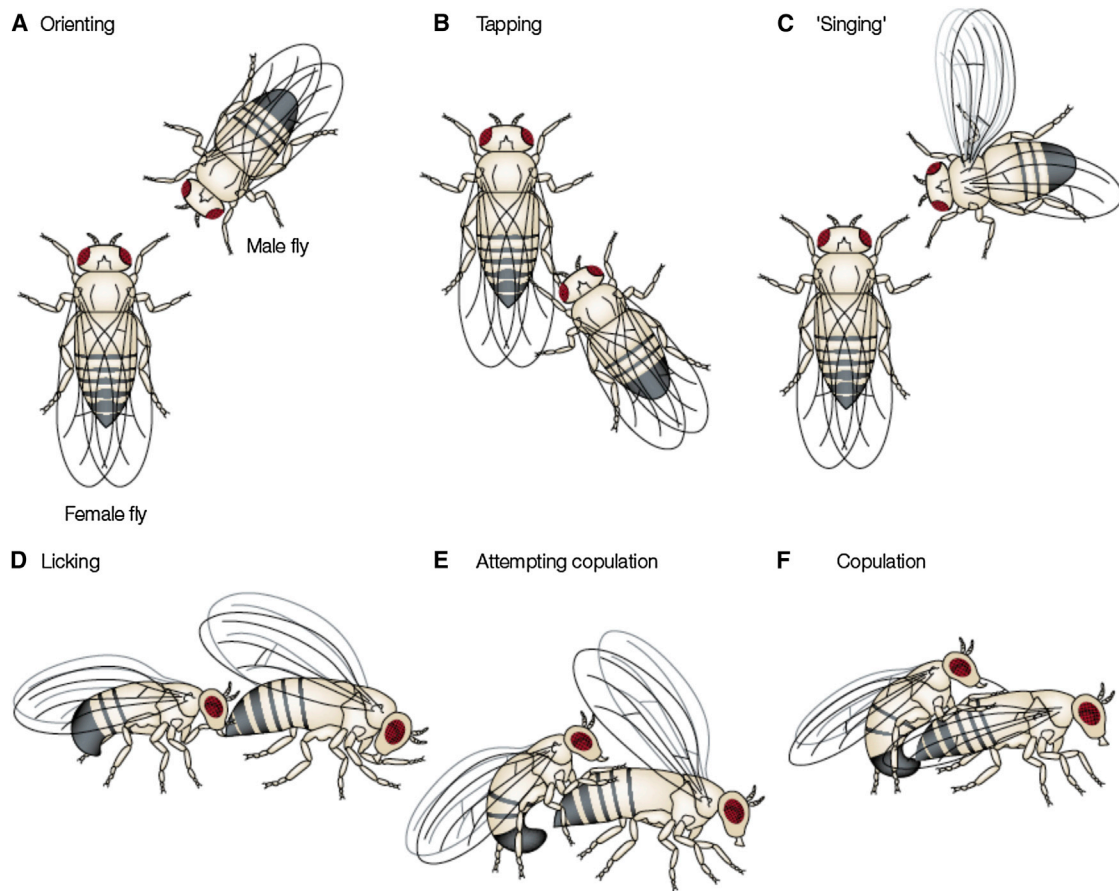


Figure 2. *Drosophila* Courtship Behavior

Sequence of human observer-defined actions in courtship behavior. The actions may vary in their duration and the length of the interval between them. The sequence is not necessarily irreversible. The “courtship index” (CI) is defined as the total amount of time that a fly spends in any of these behaviors. From Sokolowski (2001), with permission.

magnitude of this benefit. The authors used temperature-dependent manipulations of neuronal activity (Simpson, 2009) (Table 1 and Figure 1) to identify a small cluster of neurons that controls aggression. Such manipulations require a variety of controls (up to five per experiment) to exclude artifactual effects of transgenes or temperature on behavior. This, in turn, requires correction for multiple comparisons when performing statistical analysis. Because of the high variance in the data, large sample sizes (e.g., $n = 25$ per experimental or control condition) are required for sufficient statistical power (Button et al., 2013). Thus, quantifying behavior in a typical experiment with associated controls would require about 150 hr of manual scoring of aggressive behaviors (Chen et al., 2002; Dierick, 2007); the aggregate experiments in the paper would have required 2,000–3,000 hr of scoring (5 months of 40 hr/week of scoring). Such a study would, therefore, simply not have been feasible without automated measurement of behavior.

Automated analysis facilitates closed-loop experiments. Machine vision-based automated tracking of freely moving animals offers another advantage: real-time analysis of behavior. This allows on-line measurements of the animal’s position,

speed or other behavioral statistics to be fed back into a system for controlling the activity of specific neurons, in a closed-loop manner. This can reveal features of the causal relationship between neuronal activity and behavior in a manner that would not otherwise be apparent.

In a recent study, for example, online video tracking was used to target an IR laser to moving flies during courtship behavior to activate or inhibit the activity of previously identified (von Philipsborn et al., 2011) courtship neurons using thermogenetic effectors such as dTrpA1 (Hamada et al., 2008) and shibire^{ts} (Kitamoto, 2001; Bath et al., 2014) (Table 1). Using this approach, the authors were able to control neuronal activity with an unprecedented degree of spatial and temporal resolution, uncovering features of courtship behavior that had not previously been identified, such as persistent male “singing” following transient activation of a particular subset of neurons (see also Inagaki et al., 2014). This method also permits, in principle, bimodal manipulation of different neuronal populations in the fly brain, using a combination of thermogenetic and red-shifted optogenetic effectors recently adapted to flies (Lin et al., 2013; Inagaki et al., 2014; Klapoetke et al., 2014). This approach should expand

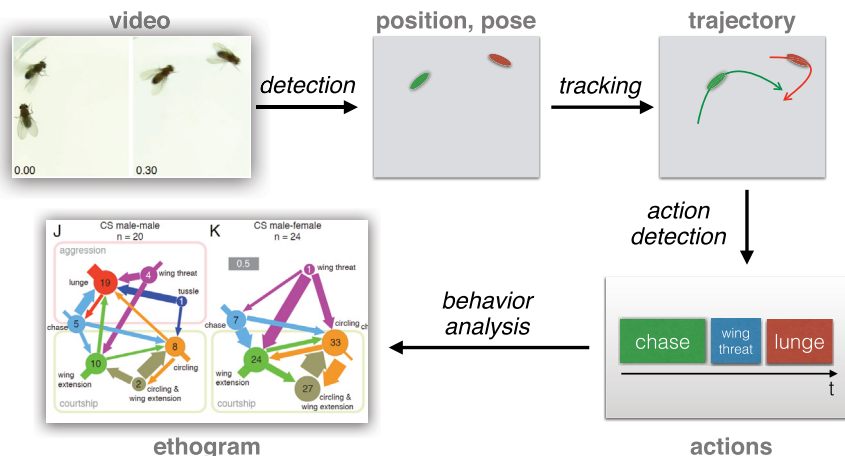


Figure 3. Summary of Steps in the Automated Analysis of Social Behavior

Each of the four steps (detection, tracking, action detection, and behavior analysis) requires validation by comparison to manually scored ground truth. The ethogram illustrates different behaviors performed during male-male and male-female social interactions. From Dankert et al., 2009, with permission. See also Table 3.

the applications of the “genetic toolkit” in *Drosophila* and facilitate the analysis of functional pathways controlling courtship and other social behaviors in *Drosophila*.

The foregoing examples all involve the quantification of behaviors that have been previously identified by biologists. The current practice is to observe a behavior and then to develop a method to automatically detect the same behavior in a manner that agrees with the human observer. However, as we discuss later, computational ethology offers the possibility of discovering new behaviors that have been missed by human observers and may soon allow us to describe and measure the complete behavioral repertoire of an animal. This in turn would enable researchers to correlate virtually any aspect of behavior with genetic, pharmacological, and neurophysiological manipulations and measurements, rather than prejudging which aspects of behavior should be measured. In order to see how this could be achieved, it is necessary to understand the process by which behavioral “classifiers” are developed using machine vision and machine learning techniques.

Anatomy of a Computational Ethology System

It has been exciting to observe, during the past decade, increasing efforts directed at automating the measurement of behavior. Both academic open-source and commercially available hardware/software packages now allow investigators to carry out measurements such as tracking the trajectories of rodents in an arena (de Chaumont et al., 2012; Gomez-Marín et al., 2012; Ohayon et al., 2013) and measuring freezing in fear conditioning experiments (Spink et al., 2001; Noldus et al., 2001; Pham et al., 2009). Such tools are valuable to the pharmaceutical industry, because they permit high-throughput screening of drugs for neurological or neuropsychiatric disorders (Tecott and Nestler, 2004). Due to the intrinsic difficulty of building automated systems, most such efforts have focused on measuring a limited but well-defined repertoire of behaviors (e.g., see Dankert et al., 2009; Pham et al., 2009). Machines or software that are created to measure a predefined set of actions, however, will be unable to reveal new ones, or new ways in which known actions may be performed. One long-term objective of computational ethology ought to be measuring an animal’s

behavior in its entirety (a discipline that has been dubbed “ethomics” [Branson et al., 2009]), including behaviors that were not foreseen by the experimentalist.

What are the components of an automated system for measuring behavior? First of all, sensors are needed. While instrumentation to detect movement (accelerometers, IR beam-breaks, reflectors), audio signals (microphones), and other physical behavioral measures have been used for some time, advances in machine vision and machine learning have led to an increased emphasis on video recording as a primary sensing modality. Cameras afford high spatial and temporal resolution and can, in principle, access most aspects of behavior without constraining the animal’s movements (Feng et al., 2004; Belongie et al., 2005; Fontaine et al., 2008). Arrangements including multiple cameras (Straw et al., 2011) and depth sensors (Shotton et al., 2013) facilitate the measurement of position and motion in 3D. Infrared cameras and lighting are used to sense the behavior of nocturnal animals, such as mice. The mundane business of acquiring and storing large volumes of video is, nevertheless, nontrivial and complicated by the need to calibrate (Hartley and Zisserman, 2003) and to synchronize all the sensors involved (Straw et al., 2011), as well as having to handle and store large volumes of data, which may reach terabytes for a single experiment (Ohayon et al., 2013).

Alongside sensors, it is often important to make use of actuators in order to modify the environment dynamically and thus elicit or modify interesting behavior; this is traditionally done by hand (Tinbergen and Perdeck, 1950). Automated and robotic systems have recently made their appearance (Reiser and Dickinson, 2008; Harvey et al., 2009; Albrecht and Bargmann, 2011), including closed-loop systems (Zabala et al., 2012; Bath et al., 2014), where the behavior of the animal determines the machine’s response, and vice-versa.

Software systems for measuring behavior from video recordings are composed of three distinct modules: tracking, action classification, and behavior analysis (Figure 3 and Table 3). Tracking means computing trajectories (Bar-Shalom, 1987). First of all, animals are detected, i.e., identified and distinguished from background in each frame of the video, and their position is measured. Additional parameters, such as the orientation of the body, limbs, and other appendages, may be measured as well; this richer characterization of the configuration of the body is often called “pose.” Finally, the position and pose of each individual animal is concatenated frame by frame to obtain a trajectory describing its motion through time (Wehrhahn et al., 1982; Noldus et al., 2001; Spink et al., 2001; Branson and

Table 3. A Glossary for Computational Ethology

Detection	Revealing the presence of an animal in a video frame and measuring its position. This requires algorithms for distinguishing the animal from the background.
Pose	The posture of an animal in three-dimensional space, including its length, orientation, and the angles of its body, limbs, or other appendages. Typically, first the animal is detected and then its pose is estimated by algorithms that analyze the image near the detected position.
Tracking	Tracing an animal's position and pose as a function of time in order to obtain a space-time trajectory from which parameters such as velocity, turning rate, and direction of approach may be computed. This requires algorithms for associating the position of an animal from one video frame to the next. Tracking is challenging when the animal moves quickly, when detection is missed in a few frames, and when the position is estimated incorrectly. These difficulties are more frequent when several animals are tracked simultaneously, since when animals interact they frequently come into contact and block each other from camera view (occlusion).
Action classifier	An algorithm using a set of rules to recognize a particular behavioral action (e.g., rearing, sniffing, wing extension) and discriminate episodes in which it occurs from those in which it does not. This includes determining the start and end points of the action (to the resolution of a single video frame). Classifiers may be hand crafted by experts or may be trained using machine learning algorithms. Classifiers are trained and tested using separate, "ground truth" data sets (see below), which contain information about tracking and pose. They can be trained in a supervised or unsupervised manner (see above).
Ground truth	Segments of video in which an animal's position, pose, and/or actions have been manually annotated, on a frame-by-frame basis, by one or more experts. Ground truth provides not only data for training classifiers, but also represent the "gold standard" for evaluating the performance of a tracker, pose estimator, or classifier.

Belongie, 2005; Straw et al., 2011; Ohayon et al., 2013; Pérez-Escudero et al., 2014). Automated tracking has been developed and applied to multiple model organisms, including *C. elegans* (Feng et al., 2004; Fontaine et al., 2006; Simonetta and Golombek, 2007; Swierczek et al., 2011), *Musca* (Wehrhahn et al., 1982), *Drosophila* (Wolf et al., 2002; Card and Dickinson, 2008; Branson et al., 2009; Dankert et al., 2009; Fontaine et al., 2009; Kohlhoff et al., 2011; Gomez-Marin et al., 2012; Iyengar et al., 2012; Tsai and Huang, 2012), ants (Khan et al., 2005), bees (Veeraraghavan et al., 2008), zebrafish (Fontaine et al., 2008; Rihel et al., 2010; Green et al., 2012), and mice (Crawley et al., 1982; Ohayon et al., 2013; Silasi et al., 2013; Weissbrod et al., 2013; Arakawa et al., 2014; Desland et al., 2014). In fact, the proliferation of automatic video tracking software has created its own problems, due to a lack of standardization. Tracking software that generalizes across multiple settings and organisms is an important step toward resolving this issue (Pérez-Escudero et al., 2014).

The next module is action classification, i.e., identifying specific intervals of time when an "action," i.e., a user-defined, relatively simple, and ethologically or ecologically relevant pattern of motion, is performed. These patterns (e.g., grooming, walking, or courtship song in flies) are detected by classifiers: computer algorithms that are able to detect instances of a given action in a video recording and to discriminate those instances from periods where the action does not occur. Action classifiers are trained from expert-provided "ground truth" examples, i.e., "positive" video clips where the action takes place and "negative" clips where it is not happening. From these labeled training examples, a machine learning algorithm can "train the classifier," i.e., develop a set of rules by which the classifier can recognize the desired actions (Dollár et al., 2005; Dankert et al., 2009; Branson et al., 2009; Jhuang et al., 2010; Burgos-Artizzu et al., 2012; de Chaumont et al., 2012; Kabra et al., 2013).

Action classifiers may be trained in two ways: "supervised" and "unsupervised." A supervised classifier is one that is trained

to detect specific actions, specified via annotated training examples (see above), that are already recognized and established by biologists, e.g., a "lunge" or a "wing extension" in *Drosophila* (Hoyer et al., 2008; Dankert et al., 2009; Iyengar et al., 2012; Tsai and Huang, 2012; Asahina et al., 2014). An unsupervised classifier is one in which the operator makes no assumptions about what kind(s) of actions and behaviors are occurring, but simply provides the learning algorithm with representative videos without annotations. Using statistical criteria, the algorithm then develops its own set of classifiers by which it decomposes the animals' behavior into units or episodes. A biologist is then free to determine whether any of those units correspond to "behaviors" that he/she could recognize. Initial efforts to apply such unsupervised approaches have recently been made in *Drosophila* (Berman et al., 2013; Vogelstein et al., 2014) (Figure 5) and in *C. elegans* (Yemini et al., 2013).

The last module, behavior analysis, has the ultimate goal of estimating "activities," i.e., large-scale behavioral patterns (aggression, courtship) assembled from different actions. At a minimum, such an analysis consists of computing an ethogram (Chen et al., 2002; Dankert et al., 2009), which describes the frequency of each action and the probability that a given action is followed by another. A more sophisticated level of analysis involves developing models of how animals make decisions and control their actions based on their internal state and on external stimuli (e.g., see Karbowski et al., 2008; Luo et al., 2010).

Each one of these modules presents considerable technical and conceptual difficulties. Tracking (of multiple animals) is made difficult by camouflage, occlusion (one animal blocking another, hiding behind an object, or leaving the field of view of the camera), motions that are too fast or too minute for the camera to follow at a given frame rate and a given spatial resolution, and trying to reconstruct three-dimensional body plans and motions based on observations made using two-dimensional sensors. These somewhat mundane difficulties may be addressed by the use of more than two cameras or of complementary

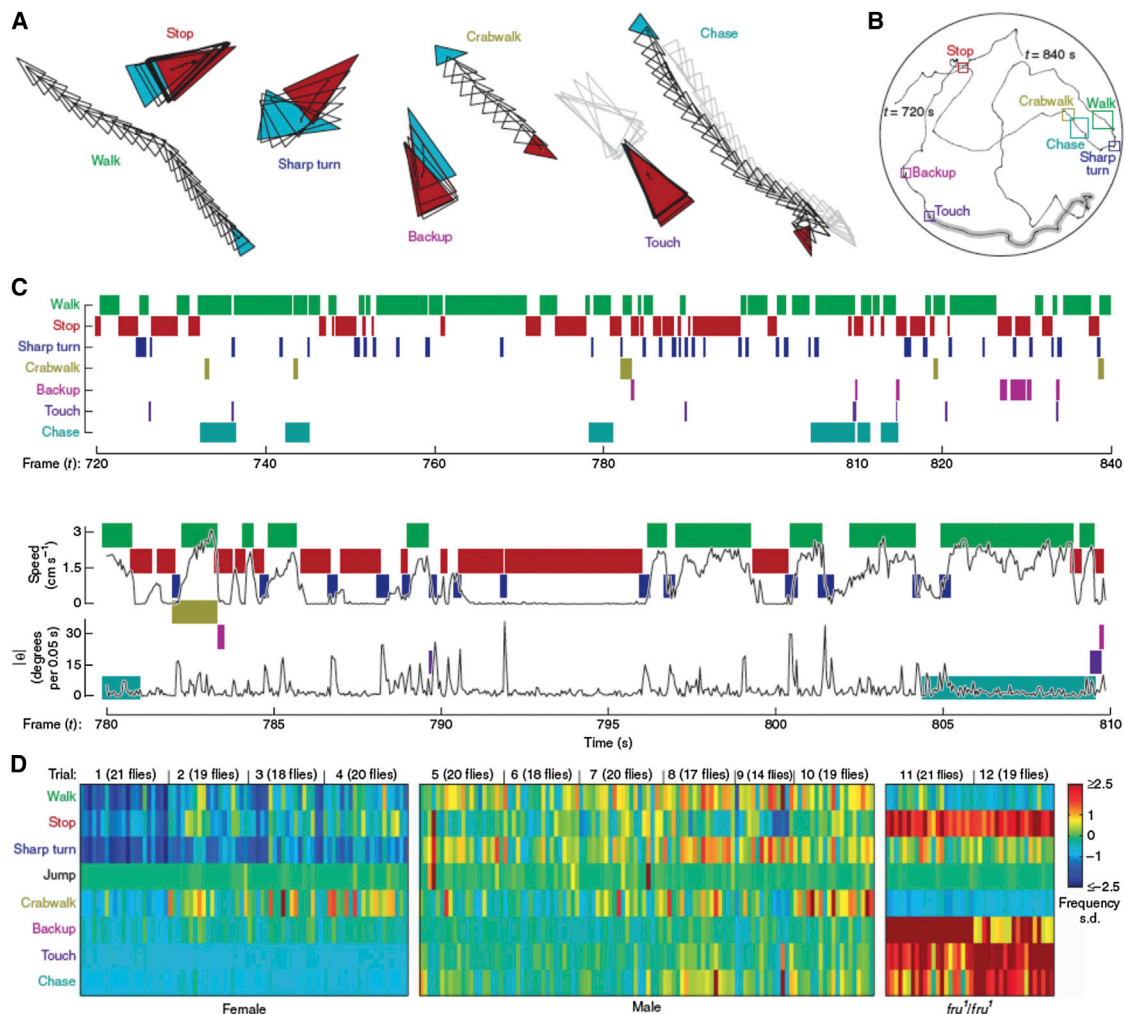


Figure 4. Ethograms Based on Machine Vision Analysis of Multiple Flies

(A–D) Eight different behaviors (A) were automatically scored from video recordings of 20 flies walking in an arena. (B) Two minute trajectory of a single male fly detected among 20 in the arena. (C) Upper: raster plot for behaviors exhibited during the trajectory in (B); lower: translational and angular velocities superimposed on a 30 s segment of the raster plot. (D) Behavioral “vectors” for female, male, and *fru1/fru1* mutant male flies. Each column represents a single fly and each row a single behavior. Numbers at top refer to experiment and number of flies tracked. From Branson et al., 2009, with permission.

technology such as depth sensors. However, there are more subtle and formidable challenges: generalization and discovery. These are currently exciting subjects of research for machine learning and machine vision researchers, and computational ethology will benefit greatly from these efforts.

Generalization

Currently, trackers, action, and behavior classifiers do not generalize well from one preparation to another (but see Pérez-Escudero et al., 2014; Kabra et al., 2013). They are designed specifically for a given animal, a given environment, and a given set of actions (Simonetta and Golombek, 2007; Huang et al., 2008; Dankert et al., 2009; Branson et al., 2009). Therefore, for each preparation a considerable investment of time and effort is required on the part of biologists, who have to label extensively by hand both the pose and the actions of a number of animals in order to provide the system with sufficient and diverse training examples. Furthermore, considerable manual effort is required

to validate and retrain the classifier until it reaches an adequate level of performance (e.g., see Asahina et al., 2014). This process must be repeated each time the system design changes or new actions are scored.

Future systems will be much more adaptable and intelligent: they will be trainable quickly and interactively to track new animals and detect new actions in a variety of experimental settings. A first step in this direction was taken in the design of the Janelia Automatic Animal Behavior Annotator (JAABA) package developed by Branson and colleagues (Kabra et al., 2013). Furthermore, by using recently developed “transfer learning” techniques (Pan and Yang, 2010), an action learned from one animal (e.g., “chasing” in *Drosophila*) will prepare the system for detecting similar actions in other animals (e.g., “chasing” in mouse).

Discovery

Human observers are sometimes able to spot novel patterns, e.g., a new behavior, or a different way in which a known action

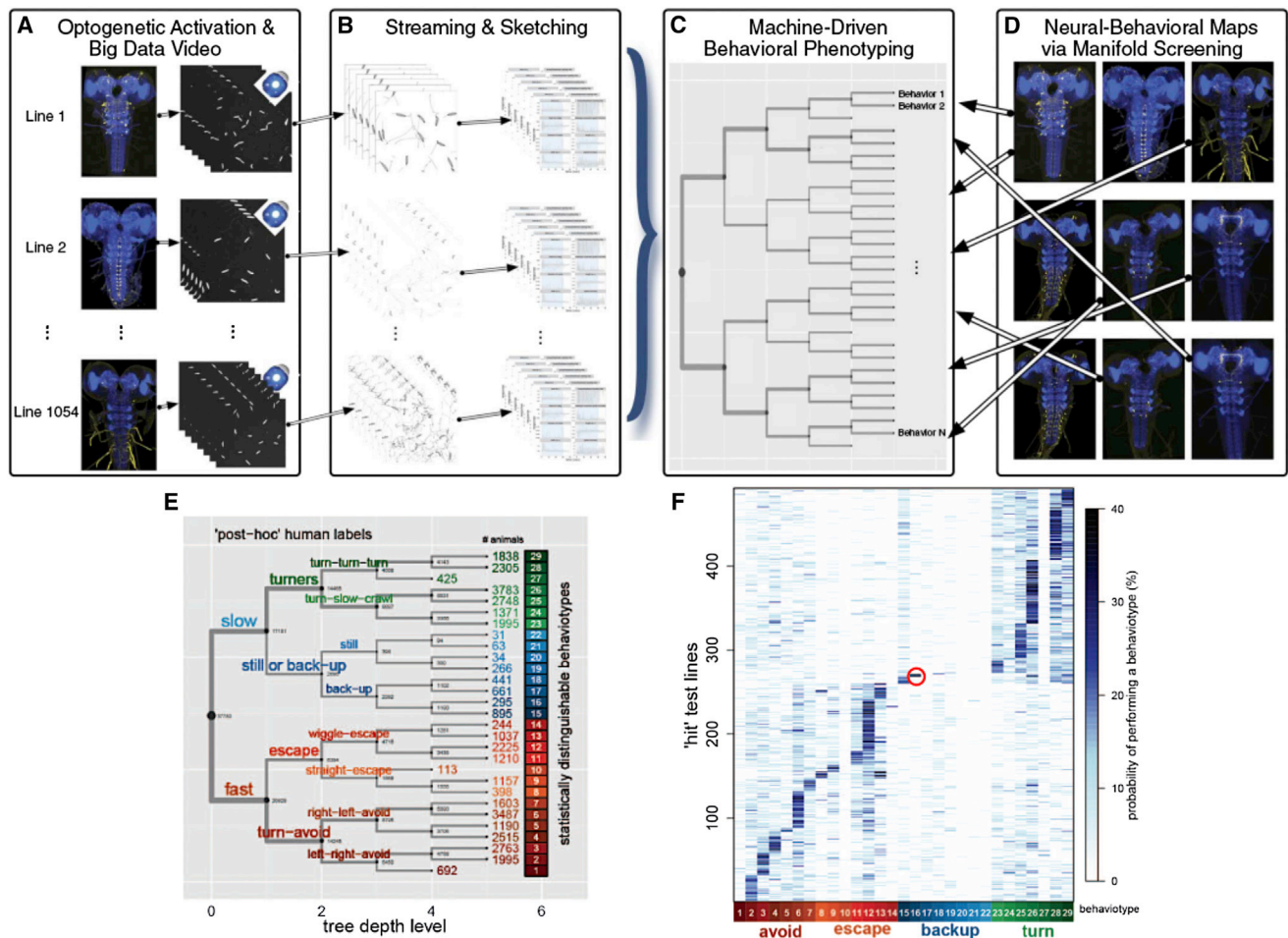


Figure 5. Unsupervised Behavioral Phenotyping of Large-Scale Neuronal Activation Screen in *Drosophila* Larvae

(A) Larvae from each of 1,050 different GAL4 lines expressing channelrhodopsin-2 in 2–15 neurons/line (37,780 animals tested in total) were photostimulated and video recorded.

(B) The contours of each larva from each video frame were extracted (streaming), and each of eight different time-varying features that characterize the shape and motion of each animal in each frame was extracted (sketching).

(C) Unsupervised machine learning was used to identify and cluster different behavioral phenotypes (“behavior types”), based on the sketches.

(D) Behavior types associated with each GAL4 line were identified.

(E) Actual tree clustering 29 different behavior types identified in the screen. Post hoc human labels were applied following retrospective analysis of videos associated with each cluster; clusters below third level could not be discriminated by human observers.

(F) Matrix illustrating the different behavior types (columns) exhibited by each of approximately 450 “hits” from the screen (GAL4 lines that were statistically significantly different from negative controls). Red circle identifies rare case of one behavior type (#16) produced by activation of just a few lines. From Vogelstein et al., 2014, with permission.

is performed, provided that they are sufficiently observant and attentive. Automated systems that have the same ability to make serendipitous discoveries would be of great value. For example, screening thousands of mutant genotypes by video recording of freely moving animals may reveal new behaviors (Brown et al., 2013; Vogelstein et al., 2014), which conventionally trained machine vision systems (i.e., supervised classifiers) are bound to either miss or to misclassify. Classifiers developed using unsupervised learning can be used to circumvent this problem and produce testable hypotheses on the existence of new behavioral phenotypes among the different genotypes screened (Brown et al., 2013; Vogelstein et al., 2014). Such techniques are reasonably well understood in the simple scenarios that are used

by theoreticians to develop proof-of-principle experiments (Gomes et al., 2010). However, their application is notoriously difficult in real-life scenarios, especially, as in the case of behavior, when the data are high-dimensional and behavioral phenomena are highly variable. A handful of pioneering studies in the past two years have started the exploration of this topic (Brown et al., 2013; Berman et al., 2013; Vogelstein et al., 2014).

A Language to Describe Behavior

How should we describe behavior? As ethology becomes more quantitative and computational, mathematical models that attempt to account for the details of phenomena appear alongside earlier qualitative verbal descriptions. We can identify three

Table 4. A Language for Behavior

Moveme	The simplest meaningful pattern associated with a behavior. Typically involves a short, ballistic trajectory described by a verb, such as a turn, a step, or a wing extension, which cannot be further decomposed. It is analogous to a “phoneme” in language.
Action	A combination of movemes that always occurs in the same stereotypical sequence and that is also described by a verb. Examples of actions include “walk” (step + step + step), “assess threat” (stop + stand up on hind legs + sniff), “eat” (open mouth + bite + chew + swallow), etc. In language, it would be analogous to a word or to an idiomatic expression.
Activity	A species-characteristic concatenation of actions and movemes whose structure is typical, ranging from stereotyped to variable. Variability can be observed both in the structure or dynamics of the individual actions that comprise an activity, as well as in the timing and/or sequence of the actions. Examples of activities include courtship, aggression, nest-building, parenting, etc.
Ethogram	A representation of the different actions that occur during an activity or activities, which indicates the frequency or probability with which each action is followed by another action (either the same or a different one). Ethograms have traditionally been computed manually, by generating a “transition matrix” composed of all the actions that are observed and the number of times each action is followed by another, given action, averaged over an observation period. Computer algorithms that detect and represent activities use hidden Markov models or other stochastic, time-series models to construct what is essentially a time-varying ethogram.
Hidden Markov model	A Markov model is a stochastic model of a time series, where knowledge of the item $X(t)$ (the “state” of the series at time t) makes the following samples independent of previous samples. For instance, if the state of the model is the action performed by a fly, the observation that a fly is lunging at time t makes any previous action irrelevant in predicting future actions (Markov property). This is clearly a simplification in most circumstances; however, biologists find Markov models informative when drawn in the form of ethograms, and engineers find them simple and useful in many circumstances. In a hidden Markov model (HMM), the state is not observable; instead, observations $Y(t)$ are dependent on the “hidden” state $X(t)$. For example, the state might be an “emotion state” of the fly (Anderson and Adolphs, 2014), and we may only be able to observe a fly’s action, which depends on the emotion. Mathematical manipulations permit estimating the hidden state of an HMM from the visible observations.

types of description/model: phenomenological, mechanistic, and functional/evolutionary. One may think of these three descriptions as answering the “what,” “how,” and “why” questions. A phenomenological description is typically the first step of analysis. However, understanding of behavior must include both a mechanistic and functional account. These are all related, but it is important to distinguish between them.

The phenomenological description accounts for the form of the phenomenon, e.g., that a “chase,” whether it is carried out by a fly or a mouse, consists of walking fast a short distance behind another individual for at least a certain amount of time. The phenomenological description is what a scientist would provide in order to characterize the action and communicate its nature to a colleague, and it is the signal that the “action classifiers” described above will use to detect and classify the action.

Since behavioral phenomena take place at multiple scales of resolution in time and space, qualitatively different statistical and geometrical models are needed to describe the phenomenology of behavior (Table 4). We use the term “moveme” (Bregler and Malik, 1998; Del Vecchio et al., 2002, 2003; Gonçalves et al., 2004) (in analogy to “phoneme”) to refer to the simplest meaningful pattern: a short, mostly ballistic, trajectory that has a verb associated with it, e.g., “turn,” “step,” “extend wing,” and cannot be further decomposed. We call “action” the composition of movemes that occur always in the same stereotypical sequence, e.g., “walk” (step + step + step), “sing” (extend wing + hold wing out + retract wing) (Siegel and Hall, 1979), and “lunge” (stop + raise on hind legs + lunge forward + stop) (Hoyer et al., 2008), and we call “activity” a concatenation of actions and movemes whose structure is typical and yet variable, e.g., “courtship” or “fighting” (Chen et al., 2002). While movemes may be detected

by frame-by-frame pattern recognition systems, detecting actions may require more sophisticated techniques to account for their composite nature (Eyjolfsson et al., 2014). Detecting and representing activities requires stochastic time-series models, such as Markov models (MM) and hidden Markov models (HMM) (Rabiner and Juang, 1986), which are commonly depicted as “ethograms” by biologists (e.g., see Yamato et al., 1992; Metz et al., 1983; Chen et al., 2002) (Table 4).

The mechanistic description relates behavior to the underlying neural mechanisms that produce it, i.e., to circuits composed of sensors, integrators, and actuators, and their relationship to the interaction of the animal’s body with the environment. For example, “chase” may be described as being triggered by visual motion, guided by visual mechanisms that respond to dark dots that, in turn, control pattern generators in the motor system (Land and Collett, 1974; Egelhaaf et al., 1988). Complex trajectories may have simple explanations once one understands the mechanisms that are involved, e.g., path integration in ants (Müller and Wehner, 1988; Collett and Collett, 2000), a bit like gravitation, provides a more satisfactory and simpler explanation of the planets’ trajectories, as observed from earth, than does Ptolemy’s theory of epicycles.

The functional/evolutionary level of description focuses on the goals or objectives of the organism under a given set of environmental and/or internal conditions. It models behavior as a means to achieve these goals, taking into account physical and information-processing constraints, but also abstracting from the specifics of the implementation. For example, the trajectory of a dragonfly chasing a prey object may be optimized to reduce the probability of its detection by the prey, and this may account for its characteristic shape (Mizutani et al., 2003).

The functional/evolutionary explanation is particularly important because it has the potential of identifying common aspects of behavior across species. This in turn can be used to formulate new questions about whether the circuit-level implementation of such common aspects is conserved or not.

The complexity of behavior makes representation and modeling a challenge. First of all, behavior is a mixture of discrete and continuous components (e.g., “decisions” and “trajectories”). Its temporal scale ranges from the simplest actions taking place in a fraction of a second (a glance, a step) to elaborate patterns lasting for minutes (courtship) or even hours and days (dominance). Furthermore, the number of individuals, objects, and locations involved in a behavior ranges in complexity from the simple act of taking a step to the multiagent, multilocation, and multiphase social behavior of a large colony of rodents.

Current modeling techniques allow us to focus on only one aspect of behavior at a time. However, as computational ethology evolves, we will seek to understand how the different components of behavior are tied together, e.g., how the instantaneous position and orientation of a fly affects the view it has of its nearest conspecific, and how this, in turn, may trigger a chase, and what effect this has on the outcome of courtship. For this type of analysis, we will need to develop more complete models of behavior that can represent all aspects of a given scene. This will require simultaneous representations at multiple timescales, discrete and continuous descriptions (actions, trajectories), multiple frames of reference (world, object-centered, partner-centered, self-centered), goals (eat, mate) as well as attributes (safe, virgin), as well as describe causal relationships. Given these complexities, it is not surprising that a general, computationally sound approach to describing behavior using conventional descriptors has not yet emerged, since it is unlikely to be manageable “by hand” (as, for instance, ethograms are). Nevertheless, important initial steps in this direction have already been taken (Berman et al., 2013; Karbowski et al., 2008; Kato et al., 2014; Luo et al., 2010). Further progress will be enabled by modern computers and software and by researchers who can seamlessly transition from biology to computation and back.

Outlook

Computational ethology offers the promise of revolutionary progress in a number of areas of neuroscience. The first is being able to combine neurophysiological recordings or optical imaging of neuronal activity, and functional manipulation of specific neurons, together with both broad and fine-grained descriptions of behavior (e.g., see Kato et al., 2014). Second, the ability to carry out rich unbiased high-throughput genetic or cellular screens will take us closer to understanding the link between genes, neural networks, and behavior (Brown et al., 2013; Vogelstein et al., 2014). These approaches will provide a wealth of detailed information allowing neuroscientists and ethologists to develop a deeper understanding of how behavior is controlled by the brain. The study of social behaviors will be particularly advanced (Branson et al., 2009; Dankert et al., 2009; Burgos-Artiz et al., 2012; Iyengar et al., 2012; Eyjolfsson et al., 2014). Ultimately, the study of human psychiatric and neurological disorders may benefit from this technology, allowing, for example, the development of novel diagnostic tests for disorders such

as autism (Hashemi et al., 2014). There are, of course, many applications of computational ethology beyond neuroscience (Coifman et al., 1998; Shotton et al., 2013; Dell et al., 2014).

As this brief (and necessarily superficial) survey indicates, CE is an expanding, fast-moving, and exciting new area of science, at the boundary between biology and engineering, where tool building, computational theory, and biological discovery are progressing hand in hand. However, many challenges remain. Behavior is among the richest and most complex phenomena that machine vision and machine learning researchers are tackling today. Progress in CE will likely become a large driver of research in these two fields. Conversely, major new insights in machine learning and machine vision are needed in order to achieve the full potential of CE in furthering neuroscience research. Lastly, and more importantly, the development of computational tools will enable researchers to study the phenomenon and causes of behavior more in detail, more completely, and more systematically.

The path ahead is exciting and rich in promise, but the endeavor is a complex intellectual and technological enterprise that will take large doses of both inspiration and perspiration to make progress and deliver on its promises. It will also require that funding agencies recognize the importance of this cutting-edge interdisciplinary field. Currently, support for research focused on the analysis of behavior per se appears as if it is being phased out, in favor of neural circuit-oriented research. While research on neural circuits is undeniably important and an exciting new frontier, our ability to pose meaningful questions about how these circuits function is, ultimately, limited by our ability to identify and measure behavioral readouts of circuit activity. Particularly in laboratory-bred model organisms, the repertoire of measurable behaviors is not especially diverse, leading to a great deal of redundant research focused on a relatively small number of behavioral paradigms. The application of CE methods, and especially unsupervised machine learning approaches, should produce a more complete and thorough characterization of known behaviors and may lead to the discovery of previously unknown behaviors, thereby diversifying the opportunities for new lines of research.

One of the biggest challenges facing CE is recruiting to the field and training a new generation of scientists who are equally conversant in machine learning and machine vision as in biology. Engineers and mathematicians are not trained to understand the nature of the questions that biologists ask, as well as the strategy that is followed by biologists to chip away at the important questions outlined above. Conversely, biologists are not routinely trained in computer science and in mathematics. Researchers in these disparate fields often do not even use compatible word-processing programs, as the coauthors of this Perspective discovered to their unending frustration. Thus, computational ethology will provide not only new scientific and technological advances, but opportunities for educating a new generation of interdisciplinary researchers as well.

AUTHOR CONTRIBUTIONS

D.J.A. and P.P. jointly shaped the message and the structure of this paper and contributed to the introduction and conclusions sections. D.J.A. contributed

writing regarding the biological applications of computational ethology. P.P. contributed writing regarding the engineering and mathematical aspects of computational ethology.

ACKNOWLEDGMENTS

The authors are grateful to the Moore Foundation, to an ONR MURI grant, and to Caltech Provost Ed Stolper for funding their research in this area. Conversations with Michael Dickinson and Kristin Branson were influential in forming many of the ideas that shaped this contribution. D.J.A. is an Investigator of the Howard Hughes Medical Institute and an Allen Distinguished Investigator. This work was supported in part by a grant from Gerald Fischbach and the Simons Foundation. D.J.A. thanks P.P. for introducing him to the joys of LaTeX.

REFERENCES

- Albrecht, D.R., and Bargmann, C.I. (2011). High-content behavioral analysis of *Caenorhabditis elegans* in precise spatiotemporal chemical environments. *Nat. Methods* 8, 599–605.
- Anderson, D.J., and Adolphs, R. (2014). A framework for studying emotions across species. *Cell* 157, 187–200.
- Anonymous. (2007). Geneticist seeks engineer: must like flies and worms. *Nat. Methods* 4, 463.
- Arakawa, T., Tanave, A., Ikeuchi, S., Takahashi, A., Kakiyama, S., Kimura, S., Sugimoto, H., Asada, N., Shiroishi, T., Tomihara, K., et al. (2014). A male-specific QTL for social interaction behavior in mice mapped with automated pattern detection by a hidden Markov model incorporated into newly developed freeware. *J. Neurosci. Methods* 234, 127–134.
- Asahina, K., Watanabe, K., Duistermars, B.J., Hoopfer, E., González, C.R., Eyjólfsson, E.A., Perona, P., and Anderson, D.J. (2014). Tachykinin-expressing neurons control male-specific aggressive arousal in *Drosophila*. *Cell* 156, 221–235.
- Bar-Shalom, Y. (1987). Tracking and data association. (San Diego: Academic Press Professional, Inc.).
- Bath, D.E., Stowers, J.R., Hörmann, D., Poehlmann, A., Dickson, B.J., and Straw, A.D. (2014). FlyMAD: rapid thermogenetic control of neuronal activity in freely walking *Drosophila*. *Nat. Methods* 11, 756–762.
- Belongie, S., Branson, K., Dollár, P., and Rabaud, V. (2005). Monitoring animal behavior in the smart vivarium. In Workshop on Measuring Behavior (The Netherlands: Wageningen).
- Bentley, D., and Konishi, M. (1978). Neural control of behavior. *Annu. Rev. Neurosci.* 1, 35–59.
- Berman, G.J., Choi, D.M., Bialek, W., and Shaveit, J.W. (2013). Mapping the stereotyped behaviour of freely-moving fruit flies. *arXiv*, arXiv:1310.4249, <http://arxiv.org/abs/1310.4249>.
- Borst, A. (2009). *Drosophila's* view on insect vision. *Curr. Biol.* 19, R36–R47.
- Branson, K., and Belongie, S. (2005). Tracking multiple mouse contours (without too many samples). In IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1, 1039–1046.
- Branson, K., Robie, A.A., Bender, J., Perona, P., and Dickinson, M.H. (2009). High-throughput ethomics in large groups of *Drosophila*. *Nat. Methods* 6, 451–457.
- Bregler, C., and Malik, J. (1998). Tracking people with twists and exponential maps. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 8–15.
- Brown, A.E., Yemini, E.I., Grundy, L.J., Jucikas, T., and Schafer, W.R. (2013). A dictionary of behavioral motifs reveals clusters of genes affecting *Caenorhabditis elegans* locomotion. *Proc. Natl. Acad. Sci. USA* 110, 791–796.
- Bui, T.V., Akay, T., Loubani, O., Hnasko, T.S., Jessell, T.M., and Brownstone, R.M. (2013). Circuits for grasping: spinal d13 interneurons mediate cutaneous control of motor behavior. *Neuron* 78, 191–204.
- Burgos-Artiz, X.P., Dollár, P., Lin, D., Anderson, D.J., and Perona, P. (2012). Social behavior recognition in continuous video. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1322–1329.
- Button, K.S., Ioannidis, J.P., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S., and Munafò, M.R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376.
- Card, G., and Dickinson, M. (2008). Performance trade-offs in the flight initiation of *Drosophila*. *J. Exp. Biol.* 211, 341–353.
- Chen, S., Lee, A.Y., Bowens, N.M., Huber, R., and Kravitz, E.A. (2002). Fighting fruit flies: a model system for the study of aggression. *Proc. Natl. Acad. Sci. USA* 99, 5664–5668.
- Coifman, B., Beymer, D., McLauchlan, P., and Malik, J. (1998). A real-time computer vision system for vehicle tracking and traffic surveillance. *Transportation Research Part C: Emerging Technologies* 6, 271–288.
- Collett, M., and Collett, T.S. (2000). How do insects use path integration for their navigation? *Biol. Cybern.* 83, 245–259.
- Crawley, J.N., Szara, S., Pryor, G.T., Creveling, C.R., and Bernard, B.K. (1982). Development and evaluation of a computer-automated color TV tracking system for automatic recording of the social and exploratory behavior of small animals. *J. Neurosci. Methods* 5, 235–247.
- Dankert, H., Wang, L., Hoopfer, E.D., Anderson, D.J., and Perona, P. (2009). Automated monitoring and analysis of social behavior in *Drosophila*. *Nat. Methods* 6, 297–303.
- de Chaumont, F., Coura, R.D., Serreau, P., Cressant, A., Chabout, J., Granon, S., and Olivo-Marin, J.-C. (2012). Computerized video analysis of social interactions in mice. *Nat. Methods* 9, 410–417, <http://dx.doi.org/10.1038/nmeth.1924>.
- Del Vecchio, D., Murray, R.M., and Perona, P. (2002). Primitives for human motion: A dynamical approach. In Proceedings of IFAC World Congress, Barcelona, Spain.
- Del Vecchio, D., Murray, R., and Perona, P. (2003). Decomposition of human motion into dynamics based primitives with application to drawing tasks. *Automatica* 39, 2085–2098.
- Dell, A.I., Bender, J.A., Branson, K., Couzin, I.D., de Polavieja, G.G., Noldus, L.P., Pérez-Escudero, A., Perona, P., Straw, A.D., Wikelski, M., and Brose, U. (2014). Automated image-based tracking and its application in ecology. *Trends Ecol. Evol.* 29, 417–428.
- Desland, F.A., Afzal, A., Warraich, Z., and Mocco, J. (2014). Manual versus Automated Rodent Behavioral Assessment: Comparing Efficacy and Ease of Bederson and Garcia Neurological Deficit Scores to an Open Field Video-Tracking System. *J. Cent. Nerv. Syst. Dis.* 6, 7–14.
- Dickinson, M.H., Farley, C.T., Full, R.J., Koehl, M.A., Kram, R., and Lehman, S. (2000). How animals move: an integrative view. *Science* 288, 100–106.
- Dickson, B.J. (2008). Wired for sex: the neurobiology of *Drosophila* mating decisions. *Science* 322, 904–909.
- Dierick, H.A. (2007). A method for quantifying aggression in male *Drosophila melanogaster*. *Nat. Protoc.* 2, 2712–2718.
- Dollár, P., Rabaud, V., Cottrell, G., and Belongie, S. (2005). Behavior recognition via sparse spatio-temporal features. In IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, pp. 65–72.
- Egelhaaf, M., Hausen, K., Reichardt, W., and Wehrhahn, C. (1988). Visual course control in flies relies on neuronal computation of object and background motion. *Trends Neurosci.* 11, 351–358.
- Eyjólfsson, E.A., Branson, S., Branson, K., Burgos-Artiz, X.P., Hoopfer, E.D., Schor, J., Anderson, D.J., and Perona, P. (2014). Detecting actions of social fruit flies. In Proceedings of the European Conference on Computer Vision (ECCV2014).
- Farrell, M.S., and Roth, B.L. (2013). Pharmacogenetics: Reimagining the pharmacogenetic approach. *Brain Res.* 1511, 6–20.

- Feng, Z., Cronin, C.J., Wittig, J.H., Jr., Sternberg, P.W., and Schafer, W.R. (2004). An imaging system for standardized quantitative analysis of *C. elegans* behavior. *BMC Bioinformatics* 5, 115.
- Fontaine, E., Burdick, J., and Barr, A. (2006). Automated tracking of multiple *C. elegans*. *Conf Proc IEEE Eng Med Biol Soc. 2006*, 3716–3719.
- Fontaine, E., Lentink, D., Kranenbarg, S., Müller, U.K., van Leeuwen, J.L., Barr, A.H., and Burdick, J.W. (2008). Automated visual tracking for studying the ontogeny of zebrafish swimming. *J. Exp. Biol.* 217, 1305–1316.
- Fontaine, E.I., Zabala, F., Dickinson, M.H., and Burdick, J.W. (2009). Wing and body motion during flight initiation in *Drosophila* revealed by automated visual tracking. *J. Exp. Biol.* 212, 1307–1323.
- Frye, M.A., and Dickinson, M.H. (2004). Closing the loop between neurobiology and flight behavior in *Drosophila*. *Curr. Opin. Neurobiol.* 14, 729–736.
- Gomes, R., Krause, A., and Perona, P. (2010). Discriminative clustering by regularized information maximization. *Advances in Neural Information Processing Systems* 23.
- Gomez-Marin, A., Partoune, N., Stephens, G.J., Louis, M., and Brembs, B. (2012). Automated tracking of animal posture and movement during exploration and sensory orientation behaviors. *PLoS ONE* 7, e41642.
- Goncalves, L., di Bernardo, E., and Perona, P. (2004). Movemes for modeling biological motion perception. *Seeing, Thinking and Knowing Theory and Decision Library A* 38, 143–170.
- Götz, K.G. (1987). Course-control, metabolism and wing interference during ultralong tethered flight in *melanogaster*. *J. Exp. Biol.* 128, 35–46.
- Green, J., Collins, C., Kyzar, E.J., Pham, M., Roth, A., Gaikwad, S., Cachat, J., Stewart, A.M., Landsman, S., Grieco, F., et al. (2012). Automated high-throughput neurophenotyping of zebrafish social behavior. *J. Neurosci. Methods* 210, 266–271.
- Hall, J.C. (1994). The mating of a fly. *Science* 264, 1702–1714.
- Hamada, F.N., Rosenzweig, M., Kang, K., Pulver, S.R., Ghezzi, A., Jegla, T.J., and Garrity, P.A. (2008). An internal thermal sensor controlling temperature preference in *Drosophila*. *Nature* 454, 217–220.
- Hamblen, M., Zehring, W.A., Kyriacou, C.P., Reddy, P., Yu, Q., Wheeler, D.A., Zwiebel, L.J., Konopka, R.J., Rosbash, M., and Hall, J.C. (1986). Germ-line transformation involving DNA from the period locus in *Drosophila melanogaster*: overlapping genomic fragments that restore circadian and ultradian rhythmicity to *per0* and *per-* mutants. *J. Neurogenet.* 3, 249–291.
- Hartley, R., and Zisserman, A. (2003). *Multiple view geometry in computer vision*. (Cambridge: Cambridge University Press).
- Harvey, C.D., Collman, F., Dombeck, D.A., and Tank, D.W. (2009). Intracellular dynamics of hippocampal place cells during virtual navigation. *Nature* 461, 941–946.
- Hashemi, J., Tepper, M., Vallin Spina, T., Esler, A., Morellas, V., Papanikolopoulos, N., Egger, H., Dawson, G., and Sapiro, G. (2014). Computer vision tools for low-cost and noninvasive measurement of autism-related behaviors in infants. *Autism Res. Treat.* 2014, 935686.
- Heisenberg, M., and Wolf, R. (1984). *Vision in: Genetics of Microbehavior*. (Berlin: Springer Verlag).
- Hoyer, S.C., Eckart, A., Herrel, A., Zars, T., Fischer, S.A., Hardie, S.L., and Heisenberg, M. (2008). Octopamine in male aggression of *Drosophila*. *Curr. Biol.* 18, 159–167.
- Huang, K.-M., Cosman, P., and Schafer, W.R. (2008). Automated detection and analysis of foraging behavior in *Caenorhabditis elegans*. *J. Neurosci. Methods* 171, 153–164.
- Inagaki, H.K., Jung, Y., Hooper, E.D., Wong, A.M., Mishra, N., Lin, J.Y., Tsien, R.Y., and Anderson, D.J. (2014). Optogenetic control of *Drosophila* using a red-shifted channelrhodopsin reveals experience-dependent influences on courtship. *Nat. Methods* 11, 325–332.
- Insel, T.R., and Winslow, J.T. (1991). Central administration of oxytocin modulates the infant rat's response to social isolation. *Eur. J. Pharmacol.* 203, 149–152.
- Iyengar, A., Imoehl, J., Ueda, A., Nirschl, J., and Wu, C.-F. (2012). Automated quantification of locomotion, social interaction, and mate preference in *Drosophila* mutants. *J. Neurogenet.* 26, 306–316.
- Jhuang, H., Garrote, E., Mutch, J., Yu, X., Khilnani, V., Poggio, T., Steele, A.D., and Serre, T. (2010). Automated home-cage behavioural phenotyping of mice. *Nat Commun* 1, 68.
- Kabra, M., Robie, A.A., Rivera-Alba, M., Branson, S., and Branson, K. (2013). JAABA: interactive machine learning for automatic annotation of animal behavior. *Nat. Methods* 10, 64–67.
- Karbowski, J., Schindelman, G., Cronin, C.J., Seah, A., and Sternberg, P.W. (2008). Systems level circuit model of *C. elegans* undulatory locomotion: mathematical modeling and molecular genetics. *J. Comput. Neurosci.* 24, 253–276.
- Kato, S., Xu, Y., Cho, C.E., Abbott, L.F., and Bargmann, C.I. (2014). Temporal responses of *C. elegans* chemosensory neurons are preserved in behavioral dynamics. *Neuron* 81, 616–628.
- Kerr, J.N., and Denk, W. (2008). Imaging in vivo: watching the brain in action. *Nat. Rev. Neurosci.* 9, 195–205.
- Khan, Z., Balch, T., and Dellaert, F. (2005). MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 1805–1819.
- Kitamoto, T. (2001). Conditional modification of behavior in *Drosophila* by targeted expression of a temperature-sensitive shibire allele in defined neurons. *J. Neurobiol.* 47, 81–92.
- Klapoetke, N.C., Murata, Y., Kim, S.S., Pulver, S.R., Birdsey-Benson, A., Cho, Y.K., Morimoto, T.K., Chuong, A.S., Carpenter, E.J., Tian, Z., et al. (2014). Independent optical excitation of distinct neural populations. *Nat. Methods* 11, 338–346.
- Kohlhoff, K.J., Jahn, T.R., Lomas, D.A., Dobson, C.M., Crowther, D.C., and Vendruscolo, M. (2011). The iFly tracking system for an automated locomotor and behavioural analysis of *Drosophila melanogaster*. *Integr. Biol.* 3, 755–760.
- Land, M.F., and Collett, T. (1974). Chasing behaviour of houseflies (*fannia canicularis*). *J. Comp. Physiol.* 89, 331–357.
- Levitin, D.A., Lidicker, W.Z., Jr., and Freund, G. (2009). Behavioural biologists don't agree on what constitutes behaviour. *Anim. Behav.* 78, 103–110.
- Lin, J.Y., Knutsen, P.M., Muller, A., Kleinfeld, D., and Tsien, R.Y. (2013). ReaChR: a red-shifted variant of channelrhodopsin enables deep transcranial optogenetic excitation. *Nat. Neurosci.* 16, 1499–1508.
- Luo, L., Callaway, E.M., and Svoboda, K. (2008). Genetic dissection of neural circuits. *Neuron* 57, 634–660.
- Luo, L., Gershow, M., Rosenzweig, M., Kang, K., Fang-Yen, C., Garrity, P.A., and Samuel, A.D. (2010). Navigational decision making in *Drosophila* thermotaxis. *J. Neurosci.* 30, 4261–4272.
- Metz, H.A., Dienes, H., de Jonge, G., and Putters, F.A. (1983). Continuous-time Markov chains as models for animal behaviour. *Bull. Math. Biol.* 45, 643–658.
- Mizutani, A., Chahl, J.S., and Srinivasan, M.V. (2003). Insect behaviour: Motion camouflage in dragonflies. *Nature* 423, 604.
- Moorman, S., Mello, C.V., and Bolhuis, J.J. (2011). From songs to synapses: molecular mechanisms of birdsong memory. Molecular mechanisms of auditory learning in songbirds involve immediate early genes, including *zenk* and *arc*, the ERK/MAPK pathway and synapsins. *Bioessays* 33, 377–385.
- Müller, M., and Wehner, R. (1988). Path integration in desert ants, *Cataglyphis fortis*. *Proc. Natl. Acad. Sci. USA* 85, 5287–5290.
- Nitabach, M.N., and Taghert, P.H. (2008). Organization of the *Drosophila* circadian control circuit. *Curr. Biol.* 18, R84–R93.
- Noldus, L.P., Spink, A.J., and Tegelenbosch, R.A. (2001). EthoVision: a versatile video tracking system for automation of behavioral experiments. *Behav. Res. Methods Instrum. Comput.* 33, 398–414.
- Ohayon, S., Avni, O., Taylor, A.L., Perona, P., and Roian Egnor, S.E. (2013). Automated multi-day tracking of marked mice for the analysis of social behaviour. *J. Neurosci. Methods* 219, 10–19.

- Pan, S.J., and Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering* 22, 1345–1359.
- Pérez-Escudero, A., Vicente-Page, J., Hinz, R.C., Arganda, S., and de Polavieja, G.G. (2014). idTracker: tracking individuals in a group by automatic identification of unmarked animals. *Nat. Methods* 11, 743–748.
- Pham, J., Cabrera, S.M., Sanchis-Segura, C., and Wood, M.A. (2009). Automated scoring of fear-related behavior using EthoVision software. *J. Neurosci. Methods* 178, 323–326.
- Rabiner, L., and Juang, B. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine* (January 1986), 4–16.
- Reichardt, W., and Poggio, T. (1976). Visual control of orientation behaviour in the fly. Part I. A quantitative analysis. *Q. Rev. Biophys.* 9, 311–375, 428–438.
- Reid, R.C. (2012). From functional architecture to functional connectomics. *Neuron* 75, 209–217.
- Reiser, M.B., and Dickinson, M.H. (2008). A modular display system for insect behavioral neuroscience. *J. Neurosci. Methods* 167, 127–139.
- Rihel, J., Prober, D.A., and Schier, A.F. (2010). Monitoring sleep and arousal in zebrafish. *Methods Cell Biol.* 100, 281–294.
- Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., Cook, M., and Moore, R. (2013). Real-time human pose recognition in parts from single depth images. *Communications of the ACM* 56, 116–124.
- Siegel, R.W., and Hall, J.C. (1979). Conditioned responses in courtship behavior of normal and mutant *Drosophila*. *Proc. Natl. Acad. Sci. USA* 76, 3430–3434.
- Silasi, G., Boyd, J.D., Ledue, J., and Murphy, T.H. (2013). Improved methods for chronic light-based motor mapping in mice: automated movement tracking with accelerometers, and chronic EEG recording in a bilateral thin-skull preparation. *Front Neural Circuits* 7, 123.
- Simonetta, S.H., and Golombek, D.A. (2007). An automated tracking system for *Caenorhabditis elegans* locomotor behavior and circadian studies application. *J. Neurosci. Methods* 161, 273–280.
- Simpson, J.H. (2009). Mapping and manipulating neural circuits in the fly brain. *Adv. Genet.* 65, 79–143.
- Sokolowski, M.B. (2001). *Drosophila*: Genetics meets behavior. *Nat. Rev. Genet.* 2, 879–890.
- Spink, A.J., Tegelenbosch, R.A., Buma, M.O., and Noldus, L.P. (2001). The EthoVision video tracking system—a tool for behavioral phenotyping of transgenic mice. *Physiol. Behav.* 73, 731–744.
- Straw, A.D., Branson, K., Neumann, T.R., and Dickinson, M.H. (2011). Multi-camera real-time three-dimensional tracking of multiple flying animals. *J. R. Soc. Interface* 8, 395–409.
- Swierczek, N.A., Giles, A.C., Rankin, C.H., and Kerr, R.A. (2011). High-throughput behavioral analysis in *C. elegans*. *Nat. Methods* 8, 592–598.
- Tataroglu, O., and Emery, P. (2014). Studying circadian rhythms in *Drosophila melanogaster*. *Methods* 68, 140–150.
- Tecott, L.H., and Nestler, E.J. (2004). Neurobehavioral assessment in the information age. *Nat. Neurosci.* 7, 462–466.
- Tinbergen, N. (1951). *The Study of Instinct*. (Oxford: Clarendon Press/Oxford University Press).
- Tinbergen, N. (1963). On aims and methods of ethology. *Z. Tierpsychol.* 20, 410–433.
- Tinbergen, N., and Perdeck, A.C. (1950). On the stimulus situation releasing the begging response in the newly hatched herring gull chick (*Larus argentatus argentatus* pont.). *Behaviour* 3, 1–39.
- Tsai, H.-Y., and Huang, Y.-W. (2012). Image tracking study on courtship behavior of *Drosophila*. *PLoS ONE* 7, e34784.
- Veeraraghavan, A., Chellappa, R., and Srinivasan, M. (2008). Shape-and-behavior encoded tracking of bee dances. *IEEE Trans. Pattern Anal. Mach. Intell.* 30, 463–476.
- Vogelstein, J.T., Park, Y., Ohya, T., Kerr, R.A., Truman, J.W., Priebe, C.E., and Zlatic, M. (2014). Discovery of brainwide neural-behavioral maps via multi-scale unsupervised structure learning. *Science* 344, 386–392.
- von Philipsborn, A.C., Liu, T., Yu, J.Y., Masser, C., Bidaye, S.S., and Dickson, B.J. (2011). Neuronal control of *Drosophila* courtship song. *Neuron* 69, 509–522.
- Wehrhahn, C., Poggio, T., and Buelthoff, H. (1982). Tracking and chasing in houseflies (musca). *Biol. Cybern.* 45, 123–130.
- Weissbrod, A., Shapiro, A., Vasserman, G., Edry, L., Dayan, M., Yitzhaky, A., Hertzberg, L., Feinerman, O., and Kimchi, T. (2013). Automated long-term tracking and social behavioural phenotyping of animal colonies within a semi-natural environment. *Nat Commun* 4, 2018.
- Williams, H. (2004). Birdsong and singing behavior. *Ann. N Y Acad. Sci.* 1016, 1–30.
- Wolf, F.W., Rodan, A.R., Tsai, L.T.-Y., and Heberlein, U. (2002). High-resolution analysis of ethanol-induced locomotor stimulation in *Drosophila*. *J. Neurosci.* 22, 11035–11044.
- Yamamoto, D., and Koganezawa, M. (2013). Genes and circuits of courtship behaviour in *Drosophila* males. *Nat. Rev. Neurosci.* 14, 681–692.
- Yamato, J., Ohya, J., and Ishii, K. (1992). Recognizing human action in time-sequential images using hidden markov model. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 379–385.
- Yemini, E., Jucikas, T., Grundy, L.J., Brown, A.E., and Schafer, W.R. (2013). A database of *Caenorhabditis elegans* behavioral phenotypes. *Nat. Methods* 10, 877–879.
- Yizhar, O., Fenno, L.E., Davidson, T.J., Mogri, M., and Deisseroth, K. (2011). Optogenetics in neural systems. *Neuron* 71, 9–34.
- Zabala, F., Polidoro, P., Robie, A., Branson, K., Perona, P., and Dickinson, M.H. (2012). A simple strategy for detecting moving objects during locomotion revealed by animal-robot interactions. *Curr. Biol.* 22, 1344–1350.