Animal-in-the-Loop: Using Interactive Robotic Conspecifics to Study Social Behavior in Animal Groups

Tim Landgraf,[1] Gregor H.W. Gebhardt,[1,2]
David Bierbach,[3,4] Pawel Romanczuk,[5] Lea Musiolek,[6]
Verena V. Hafner,[6] and Jens Krause[3,4]

[1]Department of Mathematics and Computer Science, Freie Universität Berlin, 14195 Berlin, Germany; email: tim.landgraf@fu-berlin.de

[2]Computational Systems Neuroscience, Institute of Zoology, University of Cologne, 50674 Cologne, Germany

[3]Department of Biology and Ecology of Fishes, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, 12587 Berlin, Germany

[4]Faculty of Life Sciences, Albrecht Daniel Thaer-Institute of Agricultural and Horticultural Sciences, Humboldt-Universität zu Berlin, 10099 Berlin, Germany

[5]Institute for Theoretical Biology, Department of Biology, Humboldt-Universität zu Berlin, 10115 Berlin, Germany

[6]Department of Computer Science, Humboldt-Universität zu Berlin, 10099 Berlin, Germany

Keywords

biomimetic robots, animal-in-the-loop, interactive robots, reinforcement learning

Abstract

Biomimetic robots that replace living social interaction partners can help elucidate the underlying interaction rules in animal groups. Our review focuses on the use of interactive robots that respond dynamically to animal behavior as part of a closed control loop. We discuss the most influential works to date and how they have contributed to our understanding of animal sociality. Technological advances permit the use of robots that can adapt to the situations they face and the conspecifics they encounter, or robots that learn to optimize their social performance from a set of experiences. We discuss how adaptation and learning may provide novel insights into group sociobiology and describe the technical challenges associated

with these types of interactive robots. This interdisciplinary field provides a rich set of problems to be tackled by roboticists, machine learning engineers, and control theorists. By cultivating smarter robots, we can usher in an era of more nuanced exploration of animal behavior.

1. INTRODUCTION

The use of biomimetic robots to study animal social behavior has received considerable attention in recent years. Robots that mimic the appearance and behavior of conspecifics allow biologists to embody specific hypotheses regarding social interactions and test them in the real world. Given the existence of previous reviews in this field (1–4) and the breadth of the literature, we focus here on interactive robots. These are robotic systems that are responsive to the behavior of animals— that is, the animal is part of a closed control loop, which we term animal-in-the-loop control (see **Figure 1**). Our goals are to highlight key studies that advanced the field and to analyze current trends that suggest future advances.

Incorporating the animal into the control loop is useful in two regards: It improves the social acceptance of the robot, and it increases the realism of the hybrid social dynamics, thereby allowing scientists to test hypotheses that require or relate to these dynamics. The rules by which observations are mapped to robotic actions are fixed in the majority of studies—that is, the robot is purely reactive and follows rules predefined by experts. An example of such a system is a robot that follows an animal at all times (5).
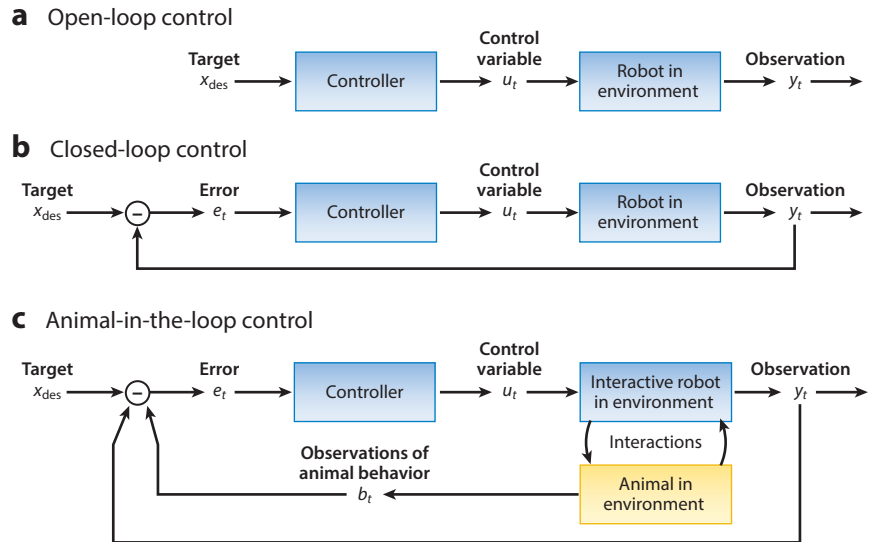

Figure 1

(*a*) Open-loop controlled robots execute their actions without any feedback from the environment. An example would be a robot that executes a specific predefined movement without taking into account other animals or obstacles in the environment. (*b*) Closed-loop controlled robots observe their environment and measure an error between the desired state of the system and the observed state. The next action is computed based on this error signal. An example could be a mobile robot that monitors its environment and avoids collisions with walls or other obstacles. (*c*) We use the term animal-in-the-loop control specifically for interactive robots that can discriminate animals from objects in the environment and consider them to be interaction partners, including behavioral observations in the error signal.

In the first half of this review, we discuss existing interactive systems. There is still much room for improvement and adoption of interactive robots. Advances in robotics and machine learning will allow the next generation of interactive robots to learn optimal behavioral responses—for example, when the robot changes its interaction rules based on prior experience. Hence, in the second part, we discuss how these new technological advances could be implemented. We focus on identifying promising concepts and discuss how the process of obtaining knowledge about the animal system may benefit from machine learning approaches.

The field of interactive robotics faces some of the same issues as other disciplines, such as computer animation (6) and virtual reality (7, 8), regarding the realism of the mimicked appearance and behavior. Much time and effort can be spent on refining the robots to create increasingly realistic interactions with animals. However, we should keep in mind that the robot and its behavior only need to be realistic enough to serve the purpose of the investigation. Most behavioral research is hypothesis driven, and the purpose of biological investigation is to have an interactive robot that allows the testing of a hypothesis. Currently, the community of behavioral robotics includes research groups that are interested primarily in pushing the capabilities of the technology as well as groups that are focused on developing new biological insights through the use of robots. In this review, we cater to both arms of the community, giving an overview of the design principles of interactive robots with reference to key studies that helped define the field. We address how interactive robots have increased our understanding of social behavior and how improved behavioral plasticity through adaptation and learning could provide novel insights into the dynamics of animal social interactions.

2. INTERACTIVE ROBOTS

Open-loop and interactive closed-loop robots have different merits for the investigation of social behavior in animals. The increasing availability of interactive robots does not mean that open-loop robots will disappear, because for many behavioral experiments, it can be useful to investigate what happens in an interaction if we switch off the responses on one side (9, 10). For example, if we are interested in the degree to which individuals vary in their responsiveness to a leader, a robot that does not respond to its social environment ensures that we are measuring individual traits uncoupled from social responses.

Established study systems that employ interactive robots include insects [cockroaches, *Periplaneta americana* (9); honeybees, *Apis mellifera* (11); and crickets, *Gryllus bimaculatus* (12)], small freshwater fish [e.g., zebrafish, *Danio rerio* (13–16); guppies, *Poecilia reticulata* (5, 17–19); weakly electric fish, *Mormyrus rume proboscirostris* (20); and golden shiners, *Notemigonus crysoleucas* (21)], rats [*Rattus norvegicus domestica* (22)], and birds [chickens, *Gallus gallus domesticus* (23, 24), and Japanese quail, *Coturnix coturnix japonica* (25)]. It is unsurprising that research has focused on small species that are easy to handle and breed in the laboratory and are already firmly embedded in the biological literature, making further research particularly relevant.

Several challenges are associated with developing interactive robots, and the specific challenge may vary drastically depending on the study species. Interactive robots require low-latency sensory feedback and enough computational power to process it in real time. While size constraints may limit the number and capabilities of sensors, actuators, and processing units, there are many examples of systems that separate the animal replica from the robot. To test hypotheses regarding the biology of animal interactions, the robot needs to reproduce relevant social cues, which in interactive robots could be responses to animal actions. While a realistic reproduction of these cues is an obvious requirement, another relevant question is whether they produce unwanted cues that could impede natural interaction, such as acoustic or vibrational noise produced by the motors

or adverse smells from nonbiological materials. Since the robot is intended to reveal information on unknown mechanisms of social interactions, the development of interactive robots is highly iterative (26). Robot development is often organized in cycles, in which we improve the robot gradually by testing various basic properties of prototypes, such as how well it is accepted by the animals. The choice of study species often limits the number of cycles. Consider, for example, a honeybee robot that mimics the waggle dance: Bees naturally use the dance when foraging for food, but they do that only in spring and summer.

The majority of interactive systems have been used in experiments that showcase how closed-loop control is beneficial for the social integration of the robot, for example, by assessing how spatial relations become more realistic as a result of integrating positional feedback information from the social environment. In the following, we give an overview of the most influential works, discussing how feedback complexity has evolved over the years and which biological questions have been addressed with this approach.

2.1. Types of Feedback Loops: What Is Investigated with Interactive Robotics?

To count as interactive, a robotic system must use at least one feedback loop with its social environment. To understand the various types of feedback loops, it is helpful to review the robot's set of sensors, which could either be carried on the robot itself (9) or be an external system [e.g., a video camera above the arena that monitors the mixed group (19)]. The simplest form of feedback can be provided with proximity sensors that detect whether a certain region close to the robot is occupied, either by a social partner or by the walls of the experimental arena.

Probably the first study to make full use of interactive technology was by Halloy et al. (9) who explored collective decision-making in cockroaches. Cockroaches and robots that smelled (but did not look) like cockroaches interacted in an experiment that explored shelter use (see **Figure 2a**). The robots were autonomous and capable of detecting the presence of cockroaches in close range. Cockroaches normally prefer the darker of two otherwise identical shelters and aggregate under it. The robots detected these locations using simple photodiodes and were programmed to behave similarly according to a mathematical model defining the probability and duration of the stay. By changing the behavioral model such that the robotic cockroaches preferred a lighter shelter, the authors showed that even a minority of robotic cockroaches were capable of steering the decision process. The mixed group now favored the lighter shelter, providing insights into the collective decision-making process that would not have been possible without robots.

Similar wheeled robots have also been used to imprint bird chicks (23–25). Young birds of these species are known to accept moving objects (regardless of their appearance) as their parents. The use of imprinting is an elegant way to get around the problem of creating realistic conspecific

Figure 2

Examples of interactive robots. (*a*) Cockroach robot. Panel adapted with permission from Reference 9. (*b*) Chicken robot (27). Photo by J. Halloy and adapted with permission. (*c*) Guppy robot (19).

robots when investigating social interactions. The PoulBot (23, 24) used both onboard sensors to detect obstacles and localize vocalizations of social partners and camera-based tracking from above the arena for rich sensory feedback (see **Figure 2b**). Gribovskiy et al. (23) proposed a method to automatically classify individual behavior over time. While the engineering efforts are impressive, most experiments conducted to date have addressed only the social cohesion of the hybrid group and have not extended into the underlying mechanisms of group dynamics.

Most systems that go beyond registering proximity employ some kind of computer vision on live camera feeds. For example, Kopman et al. (13) presented a large fish model to a single zebrafish in a separate compartment. The spatial position of the live fish in relation to the fish model was tracked with a camera from above to adjust the tailbeat frequency of the model. The authors investigated the degree to which changes in the behavior of the live fish and the model influenced each other. One drawback of the design was that the autonomous fish model had to be much larger than the live fish to accommodate the electronics and actuators, and it was therefore not recognized as a conspecific. Especially in fish studies, where the focal species may be only a few centimeters long, most studies involving interactive robots separate the control units from the animal replica to keep the robotic fish small enough to appear as a realistic conspecific. Scientists have, for example, suspended the fish replica from a rod above the tank (13, 28). The rod is then moved using a serial motion mechanism outside the tank. Another popular approach is placing the robot under the tank to move the fish replica via a magnetic coupling (17, 21, 29). A video camera above or beside the tank is used to detect and track the fish and to close the feedback loop.

In some systems, a separate camera under the tank tracks the position of the robotic unit(s) to provide accurate motion control and prevent tracking errors that mislabel live fish and robots. Landgraf et al. (18) used this approach to develop an interactive fish robot (see **Figure 2c**) that could execute a well-known shoaling model, switching to one of three behaviors [attraction, alignment, or repulsion (30)] based on the distances to the nearest neighbors. In a later study (19), the interactive robot was used to investigate different morphologies and swimming behaviors to determine which ones best achieved realistic behavioral statistics (distances to the nearest neighbor and time spent together).

Both of the above studies looked at leadership behavior but did not combine leadership and social cohesion algorithms. The latter has been explored in a recent study that investigated collective U-turns in zebrafish using three interactive robots that moved together with three live fish (31). The study compared three different robot behaviors: followership (the robot always follows the live fish), despotic (the robot moves independently of the live fish), and intermediate (the robot's directional decision depends on the perceived heading direction of its neighbors and an independent decision component). The fish robots using the intermediate algorithm were the most successful at initiating collective U-turns in the fish shoal. This approach provides an elegant way to measure the performance of the robotic algorithms in terms of successful U-turns, because U-turns are a clearly distinguishable behavior.

A different approach to study interactions between fish was taken by Worm et al. (20), who investigated the social behavior of weakly electric fish. One long-standing problem is that how electric fish adapt their pulses to particular social contexts is not fully understood. The authors moved a fish robot on predefined trajectories, but it was capable of producing pulses with millisecond latency in response to those of a live fish, which allowed the team to show that live fish likely address nearby conspecifics by echoing their signals.

Most robotic fish use two-dimensional positions of social partners as feedback and also move the robot in a fixed plane; only a few studies have made the step to controlling the robot in three-dimensional space. Kim et al. (28) presented a system that could control the three-dimensional position of a life-sized zebrafish replica. In addition, the system could imitate body-wagging

motions (minor yaw oscillations) with a fourth actuator. The authors tracked the three-dimensional position of a zebrafish and tested the effect on the fish's position in the tank with the robot following it, then compared the results with conditions in which only single axes were followed or that had an open-loop condition with predefined trajectories. This work showcases an interesting technological proof of concept and touches on the issue of how complex the models of the study systems should be.

Exploring a more complex social scenario, Shi et al. (22) developed a robot for interactions with rats. They studied the behavior of two observer rats while the rat robot was following a particular target rat. The feedback loop was closed with a camera that was positioned on top of the cage and allowed monitoring of the position of the robot in relation to the target rat that was followed and the observer rats that were avoided. The presence of the rat robot resulted in behavioral changes in the observer rats, but given the complexity of the setup, the interpretation of the results is not unambiguous.

A recent work has demonstrated an interesting variant of interactive robots. Bonnet et al. (32) presented a system that closed the loop between two species that do not interact in nature: honeybees (*Apis mellifera*) and zebrafish (*Danio rerio*). In a binary choice experiment, a system of stationary robots could affect the movement of young bees by producing heat, while the presence of bees close to the robots was measured with proximity sensors. The zebrafish robot (33) was swimming with live conspecifics in a circular corridor and was able to influence the decision of whether to swim clockwise or counterclockwise. In a closed-loop experiment, the choice of the bees for one of two robots affected the fish robot's swimming direction, and the movements of the live fish were fed back to the heating elements of the bee robots. The authors showed that both species successfully transmitted information to and used information from the partner species, showcasing that robotics can serve as a tool not only to understand intraspecies interactions but also to create and explore novel animal–robot systems.

3. ADAPTATION AND LEARNING IN INTERACTIVE ROBOTS

So far, interactive robots have produced behavior as a fixed behavioral policy $R: o \rightarrow a$ (or short policy) that maps an observation o to an action a. The policy R is predefined (usually by a human expert) and does not change during the experiment—that is, the same input o is always mapped to the same action a. Introducing behavioral plasticity through adaptation and learning can be motivated from two perspectives. From an engineering point of view, roboticists may regard the living environment as a challenging test bed for robotic systems in which learning and adaptation may be crucial for successful robotic performance. In collaborations focusing on biological questions, roboticists may want to either provide their collaborators with a robot that performs optimally (e.g., with respect to social integration) or enable research that investigates how the animals themselves learn and adapt or respond to interaction partners that learn and adapt. Virtually all animal species exhibit behavioral plasticity involving at least one of many types of learning (34, 35). From the biologist's perspective, investigating which role these biological learning processes play in animal interactions may require implementing mechanisms that allow the behavioral policy to change in response to experience. Specific hypotheses can then be tested in animal–robot interactions.

It is helpful to consider the different timescales these learning processes may span and relate them to possible classes of adaptive rules. Animals, for example, habituate to repeated stimuli with fewer responses or, by contrast, sensitize to repeated stimuli by responding more strongly (36–38). Depending on the stimulus, this type of learning can be fast: Guppies that explore a new tank may first seek the safety of a dense school and swim fast but after a few minutes will disperse more readily and slow down (10, 39, 40). However, when transferred into another unfamiliar aquarium,
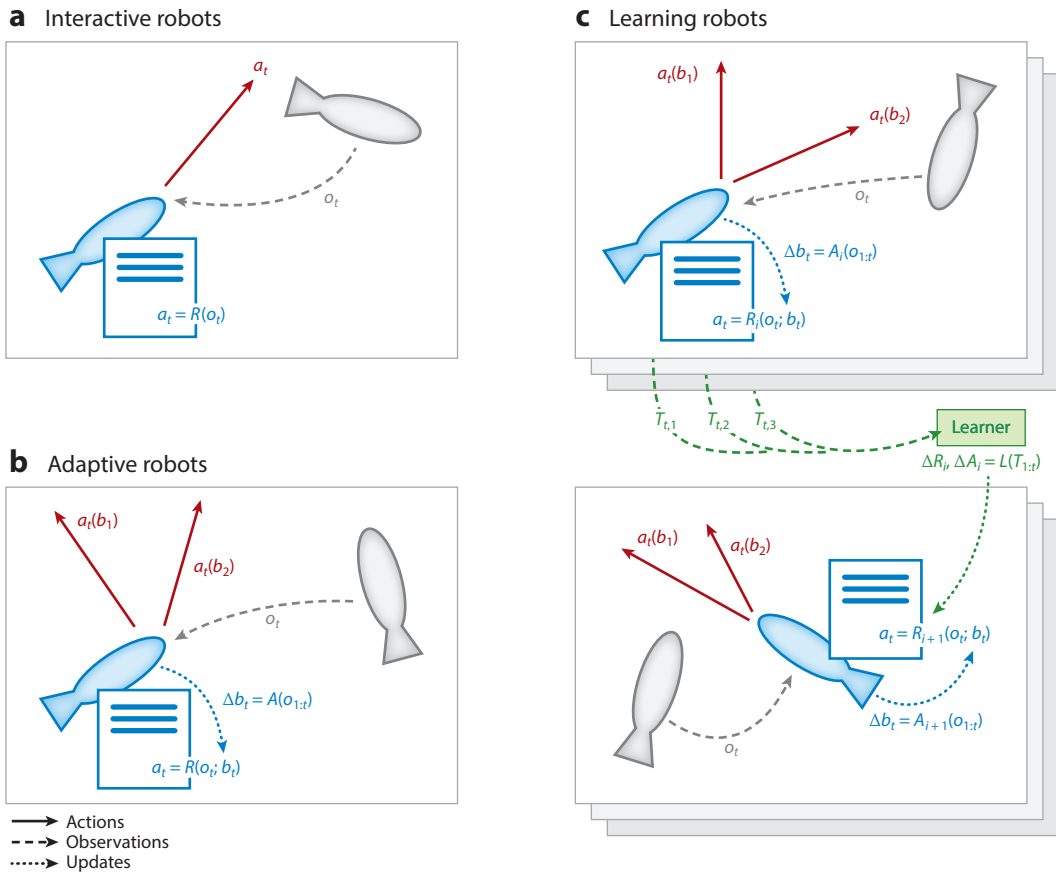
a Interactive robots

a_t

o_t

$a_t = R(o_t)$

b Adaptive robots

$a_t(b_1)$

$a_t(b_2)$

o_t

$\Delta b_t = A(o_{1:t})$

$a_t = R(o_t; b_t)$

→ Actions
⇢ Observations
⋯▸ Updates

c Learning robots

$a_t(b_1)$

$a_t(b_2)$

o_t

$\Delta b_t = A_i(o_{1:t})$

$a_t = R_i(o_t; b_t)$

$T_{t,1}$ $T_{t,2}$ $T_{t,3}$

Learner

$\Delta R_i, \Delta A_i = L(T_{1:t})$

$a_t(b_1)$ $a_t(b_2)$

$a_t = R_{i+1}(o_t; b_t)$

$\Delta b_t = A_{i+1}(o_{1:t})$

o_t

Figure 3

Interactive, adaptive, and learning robots. (*a*) Interactive robots have a fixed behavioral policy *R* that maps observations o_t of their interaction partners to actions a_t. (*b*) Adaptive robots amend the behavioral policy *R* by a parameter set *b* that they adjust during the experiment according to rules *A*. (*c*) Learning robots collect experience over many interactions and use this experience to update the behavioral policy (and in general also the adaptation rules) according to the learning algorithm *L*.

the guppies react with the same behavioral pattern. Biological learning processes underlying behavioral plasticity can also span much longer time frames. Animals may continuously integrate experience with multiple interaction partners or different environments into persistent behavioral changes [the effects of familiarity (see 41–44)]. On even longer timescales, evolution selects behavioral changes that yield higher fitness and survival [e.g., changes in sociality (45)]. Reproducing or studying short-term adaptations or long-term learning requires the robot behavior to change in response to the robot's observations. **Figure 3** illustrates the differences between purely interactive robots and adaptive/learning robots.

3.1. Adaptation Versus Learning

We define adaptive interactive robots as systems that modify the behavioral policy *R* through a set of variables *b*—that is, $a = R(o; b)$. These parameters are adjusted based on within-trial past, current, or even expected future observations according to update rules $\Delta b = A(o_{1:t})$. Adaptive robots adjust

to local social dynamics, that is, the dynamics of an individual rather than the population and of specific encounters rather than all possible situations. The adjustment of b aims to optimize a certain objective function. In contrast to machine learning, the adaptation rules may not be derived from the objective but could be designed by the experimenter. It may be unknown a priori whether a specific adaptation leads to an optimal outcome, but experiments using such an adaptation could help confirm or reject this hypothesis.

A recent study with a robotic guppy (46) provides an example of adaptive interactions. In a leadership scenario, the robot adjusts its approach vector in response to previously observed avoidance behaviors of the social partner: Every observed avoidance motion leads to an increment of a variable that in turn reduces the directness of the robot's approach, making the robot appear more careful. Long periods without avoidance behaviors, by contrast, lead to the robot's desensitization and bolder approaches. This study showed not only that this behavior decreases the amount and magnitude of the fish's avoidance behaviors but also that adaptive leaders significantly outperform leaders that do not adjust their carefulness or only adjust randomly, without regard to the observed avoidance behaviors. Adaptive leaders, for example, needed significantly fewer approaches and were followed significantly longer than nonadaptive robots. The current value of b represents a type of short-term memory of encountered past situations and adaptations. Given that the purpose of the adaptive behaviors is to track and represent local dynamics, it follows that b should disregard experiences beyond a given temporal horizon. This could, for example, relate to different experimental trials, because the animal may have been replaced, or the within-trial timeline, because the animal's internal state may have changed. These adaptive behaviors are thus quite different from purely reactive behaviors. In a study by Jolles et al. (5), a guppy robot followed individual fish, copying the forward motion speed of its partner at every point in time. Although the behavior appears to adapt to the live fish, the rules with which the robot determines its motion speed are still fixed, and the same observation will produce the same action again and again.

Learning interactive robots adjust their behavior to optimize the behavioral policy R (and in general also the adaptation rules A) for a population of situations or interaction partners. A learning rule changes the behavioral policy by $\Delta R_i = L(T_{1:t})$, where $T_{1:t}$ is the set of collected experiences up to the current time t. Typically, the learning objective or task is given in an analytical form that is amenable to numeric optimization. Adaptation and learning are not mutually exclusive, and a learned mapping may be further optimized by adaptation with a specific partner. Again, both categories may have some overlap, and a given algorithm may qualify as either adaptation or learning. We could think of an adaptive learner, for example, whose memory is large enough to contain experiences with multiple animals and is able to generalize from those experiences while simultaneously overfitting to the specific subpopulation. Learning can take place offline, using data sets obtained from the focal system, or could be executed in real time while interacting with animals in the real world or in simulations. In contrast to adaptation, learning processes aim to optimize on global social dynamics—that is, the rules are trained to exploit behavioral features that generalize well to the experienced dynamics, and potentially even new dynamics, assuming they are sufficiently well represented by the training set. In contrast to testing individual behavioral plasticity with adaptive robots, learning policies provides a means to study population-level dynamics and generic rules that apply to all data points (either multiple individuals or multiple interactions with one individual). Note that a learner may still be able to produce adaptive behavior, exploiting social rules that are observable in all animals irrespective of their individual personalities, current physiological states, or other factors that produce individual variability.

Cazenille et al. (14) presented a fish robot that automatically optimizes a set of shoaling parameters in real time. The authors measured the fitness of a given parameter configuration in episodes of 120 seconds by comparing reference distributions (natural interindividual distances, wall

HUMAN–ROBOT INTERACTION AS A SPECIAL CASE OF ANIMAL–ROBOT INTERACTION

A particular case of using robots to study the social behavior of animals is human–robot interaction. A large body of research in human–robot interaction has used both purely reactive robots and robots that adapt during the interaction. The main difference from animal studies is that interaction effects on humans are not limited to behavioral metrics: The human subjects themselves can provide valuable information through questionnaires or interviews. Even though humans may know that an artificial agent they interact with in an experiment is nonhuman, they may anthropomorphize it such that the associated brain activity is detectable and may be sufficiently aware of the anthropomorphization to report it on a questionnaire (47, 48). Indeed, in a series of correlational studies and experiments, Waytz et al. (49) showed that humans were more likely to attribute "a mind of its own" to computers and robots when the latter behaved in unpredictable ways. They explained this through the human observers' need for a predictable environment (effectance motivation): Attributing a mind to an unpredictable object made it less threatening. These results contain two relevant lessons for using interactive robots with animals: First, it may be possible to gain insights from animal–robot interactions without having to be certain that the robot is fully accepted as a conspecific, and second, it may be worth trying to manipulate the predictability of an artificial agent to see how this influences its acceptance by nonhuman animals.

Several human–robot interaction studies have provided evidence that adaptive interactions often affect human behavior much differently than purely reactive robots do, motivating similar studies in the animal–robot scenario. Schillaci et al. (50) tested different levels of robot interactiveness during a human–robot interaction experiment; using questionnaires and proximity measures, they showed that these different levels influenced user experience factors (such as excitement) and perceived robot factors (such as lifelikeness and intelligence). Scheunemann et al. (51) compared an intrinsically motivated adaptive robot with a purely reactive robot in a human–robot interaction experiment and studied differences in human perceptions of those robots. Schneider & Kummert (52) studied the effects of different levels of automation in socially assistive robots on human perception during a physical exercise scenario. In all of these studies, the difference in the level of adaptivity in the robot behavior had a strong effect on social interactions.

In their thorough review of socially aware person-following robots, Honig et al. (53) identified five layers of user needs according to which a robot may be evaluated and suggested how far factors pertaining to the human, robot, task, and environment should ideally be taken into account for each layer. This framework may exceed the level of refinement needed for research with nonhuman animals but could still be useful as a comprehensive template for evaluating a robot.

distances, and preferences for occupation of three compartments of the arena) with those measured with the robot, then performed an evolutionary optimization of the parameters for a predefined shoaling model. In this way, the robot was calibrated on a given population of 60 fishes. The paper focused on describing the impressive engineering efforts, which could then be used in future investigations of the biology of collective behavior. It would be interesting, for example, to investigate how the learning system maps different shoal compositions (think of a shoal of only bold fish and a shoal of only very shy ones) in its parameter space. It would also be interesting to see how well the learner is able to generalize from limited experience. How quickly can the system detect differences in the shoal composition?

In contrast to human–robot interaction, adaptive and learning robots have not been explored thoroughly by the animal–robot community (for additional discussion of commonalities and differences, see the sidebar titled Human–Robot Interaction as a Special Case of Animal–Robot Interaction). To stimulate further developments in the field, we next give a summary of the

components necessary for building a learning system, specify the main challenges, and discuss how learned policies may be used to augment the typical research cycle in biology.

3.2. Learning Frameworks, Action, and Observation Space

A popular machine learning approach to train behavioral policies is reinforcement learning (54). In essence, reinforcement learning is trial-and-error learning: The robot executes its current policy and collects experience in the form of actions, observations, and reward values. The reward is an expert-defined function quantifying the quality of the agent's action and the resulting state of the environment. The goal in reinforcement learning is to optimize the policy based on the experience such that the collected reward is maximized over a given time frame. For continuous action spaces, most reinforcement learning algorithms require the policy to be differentiable (e.g., neural networks). If a policy is not differentiable, then optimization methods such as evolutionary strategies (55, 56) or Bayesian optimization (57, 58) could be used. However, this generally limits the size of the search space required to converge in a feasible amount of time. The complexity of the policy model affects the required amount of data and hence the time and number of interactions that are necessary to find the optimal policy. Limiting the number of parameters could lead to models that are not complex enough to learn the required mapping from observation to optimal action. Two factors that contribute to model complexity are the choices of action and observation space.

All actions the robot can execute span an action space. The policy could, for example, control each degree of freedom of the robot, or it could just produce bearings in which the robot should move using a set of predefined motion primitives (e.g., 19). While low-level action spaces allow the policy to express a much broader range of behaviors, high-level commands help reduce the size of the search space and accelerate the learning process (59–62). This form of inductive bias allows us to restrict the search space to biologically plausible regions. From the biological perspective, movement primitives (relatively short parameterized movements that can be adapted to a certain goal and speed, such as a reaching movement of an arm) could, for example, correspond to innate behaviors of biological organisms or could be seen as basis vectors in a high-dimensional motion space from a computational perspective. These primitives could be designed by hand or result from an analysis of behavioral data (see the sidebar titled The Quantification of Behavior).

Observations are inputs to the behavioral policy and could be raw sensory data, such as video frames. Consequently, the policy model must learn to extract meaningful features and integrate how these features translate into optimal decisions. Alternatively, hand-designed features such as the relative positions of the interaction partners or the robot's simplified egocentric view of the world (as in 89), could ease the learning of the policy but again may induce human bias and reduce the expressiveness of the model. Machine learning models generally assume that the input is of fixed dimensionality. For variable environments (e.g., with a variable number of interaction partners, as in a school of fish), we need to find suitable representations. One approach could be to simplify the observations by taking only a fixed number of nearest neighbors into account when computing the action. Another widely used approach is to use a distribution over the states of the interaction partners. These distributions can then be represented by its statistics, for example, as mean and variance (90), by using mean field approximation (91), or by embedding the distribution into a reproducing kernel Hilbert space (92). Images are a widely used representation when applying deep learning methods, and using pretrained networks can reduce the amount of data and time required to learn the policy model. Other approaches to representing a variable number of observations as inputs to neural networks are deep sets (93) and deep mean embeddings (94, 95), which are network structures that process an arbitrary set of inputs and combine the output via reduction operations to a fixed-size feature vector. **Figure 4** shows examples of such observation spaces.

THE QUANTIFICATION OF BEHAVIOR

Interactive robots have strongly benefited from advances in the field of computer vision (63). Real-time video tracking often represents the essential basis of closed-loop feedback (e.g., providing the robot with the location of interaction partners). Computer vision methods are also integral to many behavioral analyses, either in guiding the development of robotic behaviors or in assessing the effect of robotic actions in the animal system.

Traditionally, animal behavior was annotated by human observers. These so-called ethograms aimed to represent an inventory of an animal's behavioral repertoire (64). In recent years, the advent of deep learning has made it possible to reliably track the position and pose of virtually any animal (65–68) and detect behaviors using hand-crafted definitions (e.g., to detect U-turns, as discussed in Section 2.1) or a machine learning model that requires training examples provided by human experts (69–71).

However, human annotations are prone to bias. Even experts may fail to detect behaviors that are relatively rare, too slow or too fast to be recognized, or too similar to be separated. Unlike supervised machine learning techniques, unsupervised methods do not rely on predefined behavioral classes and may transfer well to other species (see, e.g., Reference 72, which discovered a repertoire of 100 stereotyped behaviors in fruit flies, *Drosophila melanogaster*). Detecting distinct behavioral classes in animal behavior data may inform the design of the robot's action space and allow one to encode observations in the space of behavioral classes or even infer internal animal states (73) upon which the robot could react.

Animals may not always exhibit stereotypic (i.e., separable) behaviors, and instead may have a continuous action space. Robots could thus use time-continuous tracking data to derive meaningful information. On the individual level, typical measures of interest are movement characteristics (e.g., linear and rotational speeds and accelerations in the available dimensions) and spatiotemporal measures (e.g., time spent in a certain area of the arena or distances to the wall or other points of interest). Dyadic and group-level measures are of special interest when trying to assess the influence of a robot's behavior on live animals or vice versa. Examples include measures of whether an animal is following or avoiding another (comparing motion vectors to the respective bearings), cohesion (e.g., interindividual distances between neighbors), movement alignment (e.g., assessed by cross-correlations), and polarity (variance of the distribution of body orientations). Group dynamics can also be studied by constructing social networks and investigating their structural properties over time (74, 75).

Inferring social interactions from animal tracking data (76–78) provides important insights into possible behavioral mechanisms underlying collective behavior. However, it is often restricted by the investigation of dyads or by assumptions of specific interaction models. Furthermore, to reduce behavioral noise, these methods rely on long-time and/or large-ensemble averages, potentially ignoring individual variability and contextual, time-dependent effects (79–81).

These studies are complemented by analyses of information flows in collective animal behavior, ranging from the usage of (pairwise) time-delayed correlation functions (see, e.g., 10, 82) to information-theoretic measures such as mutual information (83), transfer entropy (84), or causation entropy (85, 86). Pilkiewicz et al. (87) provided a recent review discussing potential challenges for applying these methods in the context of collective behavior.

Empirical studies often feed into mathematical modeling, which opens up a variety of analyses—for example, regarding how sensitive the collective system is to changes of critical parameters (see, e.g., 88).

3.3. Challenges

The standard reinforcement learning setting assumes that the underlying control problem is a Markov decision process—that is, the (distribution over the) next state of the environment is fully determined by its current state and the performed action. This might be a correct assumption in robotics, where the system is fully observable, as with a robot arm that has velocity or force–torque sensors in all joints. The behavior of animals, however, is determined by unobservable or latent
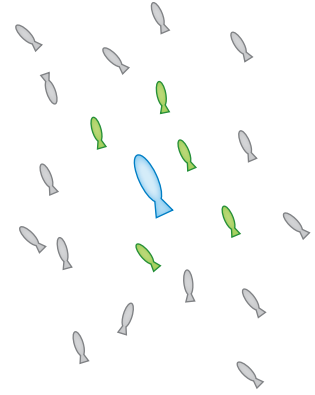
a Distance bins and raycasts

Field of view

Wall of fish tank

Blind area
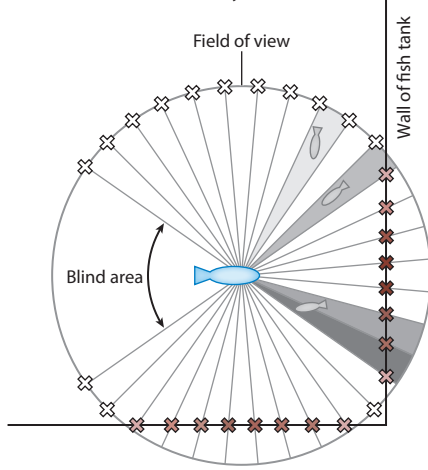
Observation as vectors:

b Distribution
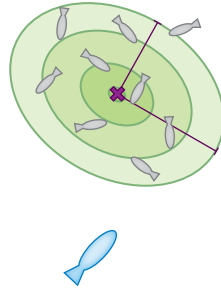
Observation as statistics,
e.g., mean and covariance:

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad \Sigma = \frac{1}{N} \sum_{i=1}^{N} (x_i - \mu)^2$$

c k nearest neighbors

Observation as $k \times m$-tuple with information
about the k nearest neighbors, e.g.,

$$(x_3, x_7, x_0, x_1, x_{10}) \quad x_i \in \mathbb{R}^m$$

Figure 4

Examples of how conspecifics can be observed by a robot. (*a*) The robot observes its interaction partners via discretized distance bins and, similarly, the distances to the walls as raycasts. The field of view is limited by a maximal observation distance and a blind area behind the robot. (*b*) A variable number of interaction partners can be represented as a distribution in which the robot observes statistics such as mean and variance. (*c*) The robot observes the *m*-dimensional state of only its *k* nearest neighbors.

variables, such as individual personality traits or past experience. For such problems, we need to assume a partially observable Markov decision process (POMDP). Here, in addition to finding the optimal policy, we need to solve the problem of inferring the latent state from past and current observations.

Since exact solutions to POMDPs are computationally intractable, practical solutions usually rely on approximations (96). As discussed above, adaptive policies are one approach to approximately solve POMDPs, since the adaptation mechanism tries to find a solution that optimizes for the latent state of the conspecific (e.g., a personality trait). Recent reinforcement learning approaches often use recurrent neural networks to model such adaptation mechanisms (97, 98). These networks have an internal state that is updated with new observations, which help to infer a latent state of the focal system. Alternatively, model-based reinforcement learning can help to solve the issues of partial observability. Instead of using adaptive policies, model-based reinforcement learning approaches learn a separate world model from observations (any observations of the focal system, even without robot interactions). This world model learns to predict the next state of the system given the current and past observations. Hence, the world model needs to maintain a representation of the system state that contains all necessary information, based on which much simpler policies can be learned. Furthermore, such world models are independent of the task for which we want to learn a policy and consequently can be reused to learn policies for different tasks.

Depending on the size of the search space, it may be prohibitive to run the learning process in real time with live animals. This problem is generally recognized in robotics, where the ability to learn control models is drastically limited by the time needed to execute a motion and the lifetime of the robot. In an animal–robot scenario, learning is further limited by the dynamic range of

social interactions, the physiological parameters of the animal, and ethical considerations. Besides simplifying the model (by, e.g., adjusting action and observation spaces) to accelerate the learning process, roboticists are often left with only one option: simulation. In virtual worlds, robots can learn the underlying rules of the environment and how the state of the environment changes after executing an action. Simulations can be run much faster than real time, and learning episodes can be run in parallel to further speed up the process. However, because simulations always differ from the real world, transferring policies from a simulation to the real world often results in suboptimal performance. How to handle this domain shift is an ongoing research question, but several approaches appear promising. Domain randomization (99) proposes to randomize the dynamics of the simulation environment so that the learner can better generalize to the dynamics in the real world; here, generalization refers to the ability to perform well on unseen data points. Zero-shot learning, one-shot learning, few-shot learning, and metalearning (100–102) are concerned with how a learner can generalize to unseen data sets with little or no extra learning.

The ability to generalize critically depends on the representativeness of the data distribution presented to the learner, whether in a simulation or in the real world. Because data sets (i.e., interaction experiences) are never complete, models cannot just learn all data by heart (i.e., we usually regularize models to limit their expressive power). However, generalizing to unseen data modes is inconceivable. How do we recognize missing modes or biases in our data? With learned rules, we must be aware that the behavior is optimal only with respect to the available data.

Let us assume that we have overcome these challenges and successfully trained a complex policy. In general, we do not explicitly know which behavioral features the policy learned to exploit or, when the training was done in a simulation, which behavioral dynamics of the simulated model are actually exploitable in the real world. Interpreting neural policies is especially hard due to the large number of parameters and activations. Animal interactions are dynamical systems best modeled with time-dependent policies, such as recurrent neural networks. How do we interpret the behavior of the policy, and what were the relevant inputs? Complex recurrent architectures may use several layers of hidden states that represent the agent's history. What do these states actually mean? In recent years, the field of explainable AI has generated several approaches to analyze, visualize, and explain the way neural networks map inputs to outputs (103). Especially in the domain of image recognition tasks, a variety of approaches have been proposed; however, most successful work has focused primarily on convolutional neural networks, a special type of feed-forward network.

3.4. Use Cases for Adaptation and Learning in Animal–Robot Interactions

Despite the challenges, learning and adaptation offer several interesting use cases in the animal-robot interaction scenario:

1. Understanding behavior as a result of multiple drivers under multiple constraints: Learning under different constraints or using different reward functions will yield different policies. Studying how they differ behaviorally (e.g., whether they express biologically plausible behavior) may provide insights into which selection pressures contributed to the evolution of natural behavior.

2. Using learning robots as a tool for automatic phenotyping: Animals of the same species often differ consistently in their behavior (104), and even clonal animals reared in the same environment show variation in their behavioral types (105, 106). Studying these differences often requires biologists to repeatedly conduct multiple behavioral assays (e.g., running a specific experimental protocol), each used to assess an animal's personality along a specific dimension. While open- and closed-loop robots are already used to provide animals with

standardized social cues and thus work out the live fish's individual differences (5, 10), adaptive robots could be used to automate this process and steer the animal to behavioral dynamics that are more informative than would be possible without adaptation. The way the robot adapts, or the configuration of parameters it converges on, may be predictive of a specific behavioral phenotype.

3. Evaluating animal behavior models: Let P_A and P_B be two policies trained on animal interaction models A and B, respectively, to maximize the same objective function $\mathcal{J}$ in simulation. When A and B are representing similar but different dynamics, one policy may learn to exploit interaction dynamics unavailable to the other to maximize $\mathcal{J}$. Since A and B are only incompletely representing the real dynamics, both policies will likely perform worse in the real world than they did in simulation. The dynamics the policies have learned to exploit (or a subset thereof) may still be available to the robot. Testing and comparing robot-embedded policies may then reveal which model is more realistic regarding the task represented by $\mathcal{J}$, which at the same time serves as an evaluation metric in the real world. This way, animal interaction hypotheses compete in the same frame of reference, which yields a fairer model comparison than, for instance, defining metrics that are independent of task and therefore could favor one model over another.

4. Optimizing the environment for hypothesis testing: The laboratory environments in which we run experiments are tuned to test specific hypotheses. Learning can be used to optimize not only robotic behavior but also any architectural structure or dynamical cue to stimulate the animals (e.g., light, sound, and the geometry of the arena). However, most high-dimensional search spaces may have to resort to simulated dynamics, and without the data that can represent the real dynamics, we find ourselves in a dilemma: How do we simulate the dynamics when we do not know what they are in the real world?

5. Enabling curiosity-based exploration to learn policies that elicit new data: Model-based approaches may represent a way out of the dilemma mentioned above (107–109). The behavioral policy, trained using interactions with a world model instead of the real world, could use curiosity or intrinsic motivation (110–114) to explore the environment in dynamics regions that are poorly represented by the model because data were missing. The model can then be improved by iteratively finding poorly supported areas in the learned world model and collecting new data from interactions in the real environment. These regions of the social dynamics may, however, be hard to sample due to local instabilities of the system. This may motivate close connections to mathematical modeling to study these systems (e.g., with dynamical system analysis).

4. CONCLUSIONS

Interactive robots are still a technological challenge and by no means trivial to operate. Therefore, many studies are still engaged in the important task of improving the capability of these systems rather than homing in on answering biological questions of significance. Progress in the field of interactive robots clearly requires both further developments in robotics and a clear identification of important and relevant research questions. Here, we identify three areas that we believe are of high biological significance.

The first is leadership. Large parts of the literature on collective behavior and collective intelligence remain untested, because simulations of multiagent interactions can be carried out much faster than empirical tests, which means that a large body of untested or only partly tested theory exists. Simulations have proposed solutions to leadership strategies as a function of group size (115). For example, an individual that wants to lead one or more conspecifics to a target location is

faced with making a trade-off between goal-oriented movement and social cohesion. If too much emphasis is put on goal-oriented behavior, then groups tend to split, but if goal-oriented movement is too weak, then it takes a long time to reach the target. Ioannou et al. (116) carried out some elegant experiments on golden shiners (*Notemigonus crysoleucas*) that used trained fish that differed in their goal-oriented movement in social interactions. In the future, trained fish could be substituted with robotic fish, allowing investigators to specifically control the balance between goal-oriented behavior and social cohesion and vary it to determine when each type of leadership is most critical. Testing leadership as a function of group size, however, is only one interesting option. Tests that experimentally manipulate the composition of groups would also be of great interest, and adaptive interactive robots allow us to examine the role of social responsiveness in such processes.

Animal cognition is a second significant area. Frohnwieser et al. (117) argued that robots could help address fundamental questions in comparative psychology regarding perception, spatial cognition, social cognition, and early cognitive development. Interactive robots have great potential when it comes to the questions of whether and (if so) to what degree animals can anticipate and predict the behavior of con- and heterospecifics. A closely connected question is whether the ability to predict the behavior of others is used to influence the outcome of interactions. The highly controlled nature of social interactions involving robots makes it possible to address these questions in ways that were not possible before. Dyadic interactions in which a robot and a biological agent try to outsmart each other could be a promising testing ground for future experiments, providing insights into the strategic capabilities of animal minds.

Finally, heterospecific interactions and group hunting are receiving increased interest (118). We have a fairly good understanding of some aspects of the collective behavior of conspecific animals, but little is known about the interactions between social predators and group-living prey or about their potential coevolution. Robots can potentially help us to better understand the strengths and weaknesses of collective prey behavior when under attack by cooperatively hunting predators. Robotic models of predators can be made to interact according to particular algorithms to study the effects on the collective behavior of prey.

Machine learning has been applied to a variety of problem settings. We focused in this review on animal social interactions, but promising advances have also been made in other related fields. Bieker et al. (119) used a deep recurrent neural network to model complex fluid dynamics and learn a controller that could successfully force the system to a desired state. Analogously, neuroscientists have shown that it is possible to simultaneously record from and stimulate brain tissue, and to use machine learning to drive neural activity to desired states (120) and characterize computational properties of the tissue (121).

Despite its growing popularity, the use of robots to investigate social behavior in animals is still at its beginning. Many behavioral biologists who could benefit enormously from testing their hypotheses with interactive robots still find the interdisciplinary threshold too high to engage with roboticists. At the same time, many roboticists working on animals could benefit from the conceptual input of biologists and the interesting problem settings in complex animal systems. These two sides have much to gain from each other, and we hope that this review will be instrumental in bringing them closer together to generate more exciting research that pushes the boundaries of both robotics and behavioral biology.

DISCLOSURE STATEMENT

ACKNOWLEDGMENTS

LITERATURE CITED

1. Webb B. 2000. What does robotics offer animal behaviour? *Anim. Behav.* 60:545–58
2. Krause J, Winfield AF, Deneubourg JL. 2011. Interactive robots in experimental biology. *Trends Ecol. Evol.* 26:369–75
3. Romano D, Donati E, Benelli G, Stefanini C. 2018. A review on animal–robot interaction: from bio-hybrid organisms to mixed societies. *Biol. Cybernet.* 113:201–25
4. Datteri E. 2020. Interactive biorobotics. *Synthese.* In press. **https://doi.org/10.1007/s11229-020-02533-2**
5. Jolles JW, Weimar N, Landgraf T, Romanczuk P, Krause J, Bierbach D. 2020. Group-level patterns emerge from individual speed as revealed by an extremely social robotic fish. *Biol. Lett.* 16:20200436
6. Chouinard-Thuly L, Gierszewski S, Rosenthal GG, Reader SM, Rieucau G, et al. 2017. Technical and conceptual considerations for using animated stimuli in studies of animal behavior. *Curr. Zool.* 63:5–19
7. Moussaïd M, Kapadia M, Thrash T, Sumner RW, Gross M, et al. 2016. Crowd behaviour during high-stress evacuations in an immersive virtual environment. *J. R. Soc. Interface* 13:20160414
8. Stowers JR, Hofbauer M, Bastien R, Griessner J, Higgins P, et al. 2017. Virtual reality for freely moving animals. *Nat. Methods* 14:995–1002
9. Halloy J, Sempo G, Caprari G, Rivault C, Asadpour M, et al. 2007. Social integration of robots into groups of cockroaches to control self-organized choices. *Science* 318:1155–58
10. Bierbach D, Landgraf T, Romanczuk P, Lukas J, Nguyen H, et al. 2018. Using a robotic fish to investigate individual differences in social responsiveness in the guppy. *R. Soc. Open Sci.* 5:181026
11. Landgraf T, Oertel M, Rhiel D, Rojas R. 2010. A biomimetic honeybee robot for the analysis of the honeybee dance communication system. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 3097–102. Piscataway, NJ: IEEE
12. Kawabata K, Aonuma H, Hosoda K, Sugimoto Y, Xue J. 2014. Experimental study on robotic interactions to the cricket. In *2014 IEEE International Conference on Robotics and Biomimetics*, pp. 949–54. Piscataway, NJ: IEEE
13. Kopman V, Laut J, Polverino G, Porfiri M. 2013. Closed-loop control of zebrafish response using a bioinspired robotic-fish in a preference test. *J. R. Soc. Interface* 10:20120540
14. Cazenille L, Chemtob Y, Bonnet F, Gribovskiy A, Mondada F, et al. 2018. How to blend a robot within a group of zebrafish: achieving social acceptance through real-time calibration of a multi-level behavioural model. In *Biomimetic and Biohybrid Systems*, ed. V Vouloutsi, J Halloy, A Mura, M Mangan, N Lepora, et al., pp. 73–84. Cham, Switz.: Springer
15. Spinello C, Yang Y, Macr S, Porfiri M. 2019. Zebrafish adjust their behavior in response to an interactive robotic predator. *Front. Robot. AI* 6:38
16. Chemtob Y, Cazenille L, Bonnet F, Gribovskiy A, Mondada F, Halloy J. 2020. Strategies to modulate zebrafish collective dynamics with a closed-loop biomimetic robotic system. *Bioinspir. Biomimet.* 15:046004
17. Landgraf T, Nguyen H, Forgo S, Schneider J, Schröer J, et al. 2013. Interactive robotic fish for the analysis of swarm behavior. In *Advances in Swarm Intelligence*, ed. Y Tan, Y Shi, H Mo, pp. 1–10. Berlin: Springer
18. Landgraf T, Nguyen H, Schröer J, Szengel A, Clément RJG, et al. 2014. Blending in with the shoal: robotic fish swarms for investigating strategies of group formation in guppies. In *Biomimetic and Biohybrid Systems*, ed. A Duff, NF Lepora, A Mura, TJ Prescott, PFMJ Verschure, pp. 178–189. Cham, Switz.: Springer

19. Landgraf T, Bierbach D, Nguyen H, Muggelberg N, Romanczuk P, Krause J. 2016. RoboFish: increased acceptance of interactive robotic fish with realistic eyes and natural motion patterns by live Trinidadian guppies. *Bioinspir. Biomimet.* 11:015001

20. Worm M, Landgraf T, Prume J, Nguyen H, Kirschbaum F, von der Emde G. 2018. Evidence for mutual allocation of social attention through interactive signaling in a mormyrid weakly electric fish. *PNAS* 115:6852–57

21. Swain DT, Couzin ID, Leonard NE. 2012. Real-time feedback-controlled robotic fish for behavioral experiments with fish schools. *Proc. IEEE* 100:150–63

22. Shi Q, Ishii H, Tanaka K, Sugahara Y, Takanishi A, et al. 2015. Behavior modulation of rats to a robotic rat in multi-rat interaction. *Bioinspir. Biomimet.* 10:056011

23. Gribovskiy A, Mondada F, Deneubourg JL, Cazenille L, Bredeche N, Halloy J. 2015. Automated analysis of behavioural variability and filial imprinting of chicks (*G. gallus*), using autonomous robots. arXiv:1509.01957 [q-bio.QM]

24. Gribovskiy A, Halloy J, Deneubourg JL, Mondada F. 2018. Designing a socially integrated mobile robot for ethological research. *Robot. Auton. Syst.* 103:42–55

25. Jolly L, Pittet F, Caudal JP, Mouret JB, Houdelier C, et al. 2016. Animal-to-robot social attachment: initial requisites in a gallinaceous bird. *Bioinspir. Biomimet.* 11:016007

26. Mondada F, Martinoli A, Correll N, Gribovskiy A, Halloy JI, et al. 2013. A general methodology for the control of mixed natural-artificial societies. In *Handbook of Collective Robotics*, ed. S Kernbach, pp. 547–86. Singapore: Pan Stanford

27. Gribovskiy A, Halloy J, Deneubourg JL, Bleuler H, Mondada F. 2010. Towards mixed societies of chickens and robots. In *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4722–28. Piscataway, NJ: IEEE

28. Kim C, Ruberto T, Phamduy P, Porfiri M. 2018. Closed-loop control of zebrafish behaviour in three dimensions using a robotic stimulus. *Sci. Rep.* 8:657

29. Bonnet F, Binder S, de Oliveria ME, Halloy J, Mondada F. 2014. A miniature mobile robot developed to be socially integrated with species of small fish. In *2014 IEEE International Conference on Robotics and Biomimetics*, pp. 747–52. Piscataway, NJ: IEEE

30. Couzin ID, Krause J, James R, Ruxton GD, Franks NR. 2002. Collective memory and spatial sorting in animal groups. *J. Theor. Biol.* 218:1–11

31. Papaspyros V, Bonnet F, Collignon B, Mondada F. 2019. Bidirectional interactions facilitate the integration of a robot into a shoal of zebrafish *Danio rerio*. *PLOS ONE* 14:e0220559

32. Bonnet F, Mills R, Szopek M, Schönwetter-Fuchs S, Halloy J, et al. 2019. Robots mediating interactions between animals for interspecies collective behaviors. *Sci. Robot.* 4:eaau7897

33. Bonnet F, Gribovskiy A, Halloy J, Mondada F. 2018. Closed-loop interactions between a shoal of zebrafish and a group of robotic fish in a circular corridor. *Swarm Intell.* 12:227–44

34. Mery F, Burns JG. 2010. Behavioural plasticity: an interaction between evolution and experience. *Evol. Ecol.* 24:571–83

35. Snell-Rood EC. 2013. An overview of the evolutionary causes and consequences of behavioural plasticity. *Anim. Behav.* 85:1004–11

36. Rankin CH, Abrams T, Barry RJ, Bhatnagar S, Clayton DF, et al. 2009. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiol. Learn. Mem.* 92:135–38

37. Peeke HVS, Herz MJ, eds. 2012. *Habituation*, Vol. 1: *Behavioral Studies*. New York: Academic

38. Blumstein DT. 2016. Habituation and sensitization: new thoughts about old ideas. *Anim. Behav.* 120:255–62

39. Mikheev VN, Andreev OA. 1993. Two-phase exploration of a novel environment in the guppy, *Poecilia reticulata*. *J. Fish Biol.* 42:375–83

40. Lukas J, Kalinkat G, Miesen FW, Landgraf T, Krause J, Bierbach D. 2020. Consistent behavioral syndromes across seasons in an invasive freshwater fish. bioRxiv 2020.03.03.974998. **https://doi.org/10.1101/2020.03.03.974998**

41. Kelley JL, Graves JA, Magurran AE. 1999. Familiarity breeds contempt in guppies. *Nature* 401:661–62

42. Brown C. 2001. Familiarity with the test environment improves escape responses in the crimson spotted rainbowfish, *Melanotaenia duboulayi*. *Anim. Cogn.* 4:109–13

43. Bierbach D, Girndt A, Hamfler S, Klein M, Mücksch F, et al. 2011. Male fish use prior knowledge about rivals to adjust their mate choice. *Biol. Lett.* 7:349–51

44. Doran C, Bierbach D, Laskowski KL. 2019. Familiarity increases aggressiveness among clonal fish. *Anim. Behav.* 148:153–59

45. Bierbach D, Krause S, Romanczuk P, Lukas J, Arias-Rodriguez L, Krause J. 2020. An interaction mechanism for the maintenance of fission–fusion dynamics under different individual densities. *PeerJ* 8:e8974

46. Landgraf T, Moenck HJ, Gebhardt GHW, Weimar N, Hocke M, et al. 2020. Socially competent robots: adaptation improves leadership performance in groups of live fish. arXiv:2009.06633 [cs.RO]

47. Harris LT, Fiske ST. 2008. The brooms in Fantasia: neural correlates of anthropomorphizing objects. *Soc. Cogn.* 26:210–23

48. Krach S, Hegel F, Wrede B, Sagerer G, Binkofski F, Kircher T. 2008. Can machines think? Interaction and perspective taking with robots investigated via fMRI. *PLOS ONE* 3:e2597

49. Waytz A, Morewedge CK, Epley N, Monteleone G, Gao JH, Cacioppo JT. 2010. Making sense by making sentient: effectance motivation increases anthropomorphism. *J. Pers. Soc. Psychol.* 99:410–35

50. Schillaci G, Bodiroža S, Hafner VV. 2013. Evaluating the effect of saliency detection and attention manipulation in human-robot interaction. *Int. J. Soc. Robot.* 5:139–52

51. Scheunemann MM, Salge C, Dautenhahn K. 2019. Intrinsically motivated autonomy in human-robot interaction: human perception of predictive information in robots. In *Towards Autonomous Robotic Systems*, ed. K Althoefer, J Konstantinova, K Zhang, pp. 325–37. Cham, Switz.: Springer

52. Schneider S, Kummert F. 2020. Comparing robot and human guided personalization: Adaptive exercise robots are perceived as more competent and trustworthy. *Int. J. Soc. Robot.* In press. **https://doi.org/10.1007/s12369-020-00629-w**

53. Honig SS, Oron-Gilad T, Zaichyk H, Sarne-Fleischmann V, Olatunji S, Edan Y. 2018. Toward socially aware person-following robots. *IEEE Trans. Cogn. Dev. Syst.* 10:936–54

54. Sutton RS, Barto AG. 2018. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press

55. Rechenberg I. 1973. *Evolutionsstrategie, Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Stuttgart, Ger.: Frommann-Holzboog

56. Hansen N, Ostermeier A. 2001. Completely derandomized self-adaptation in evolution strategies. *Evol. Comput.* 9:159–95

57. Lizotte D, Wang T, Bowling M, Schuurmans D. 2007. Automatic gait optimization with Gaussian process regression. In *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, pp. 944–49. San Francisco: Morgan Kaufmann

58. Wilson A, Fern A, Tadepalli P. 2010. Incorporating domain models into Bayesian optimization for RL. In *Machine Learning and Knowledge Discovery in Databases*, ed. JL Balcázar, F Bonchi, A Gionis, M Sebag, pp. 467–82. Berlin: Springer

59. Stolle M, Precup D. 2002. Learning options in reinforcement learning. In *International Symposium on Abstraction, Reformulation, and Approximation*, ed. S Koenig, RC Holte, pp. 212–23. Berlin: Springer

60. Sutton RS, Precup D, Singh S. 1999. Between MDPs and semi-MDPs: a framework for temporal abstraction in reinforcement learning. *Artif. Intell.* 112:181–211

61. Schaal S, Peters J, Nakanishi J, Ijspeert A. 2005. Learning movement primitives. In *Robotics Research: The Eleventh International Symposium*, ed. B Siciliano, O Khatib, P Dario, R Chatila, pp. 561–72, Berlin: Springer

62. Ude A, Gams A, Asfour T, Morimoto J. 2010. Task-specific generalization of discrete and periodic dynamic movement primitives. *IEEE Trans. Robot.* 26:800–15

63. Dell AI, Bender JA, Branson K, Couzin ID, de Polavieja GG, et al. 2014. Automated image-based tracking and its application in ecology. *Trends Ecol. Evol.* 29:417–28

64. Tinbergen N. 1951. *The Study of Instinct*. New York: Clarendon

65. Romero-Ferrero F, Bergomi MG, Hinz RC, Heras FJH, de Polavieja GG. 2019. idtracker.ai: tracking all individuals in small or large collectives of unmarked animals. *Nat. Methods* 16:179–82

66. Pereira TD, Aldarondo DE, Willmore L, Kislin M, Wang SSH, et al. 2019. Fast animal pose estimation using deep neural networks. *Nat. Methods* 16:117–25

67. Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, et al. 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* 21:1281–89

68. Graving JM, Chae D, Naik H, Li L, Koger B, et al. 2019. DeepPoseKit, a software toolkit for fast and robust animal pose estimation using deep learning. *eLife* 8:e47994

69. Dankert H, Wang L, Hoopfer ED, Anderson DJ, Perona P. 2009. Automated monitoring and analysis of social behavior in *Drosophila*. *Nat. Methods* 6:297–303

70. Branson K, Robie A, Bender J, Perona P, Dickinson M. 2009. High-throughput ethomics in large groups of *Drosophila*. *Nat. Methods* 6:451–57

71. Kain J, Stokes C, Gaudry Q, Song X, Foley J, et al. 2013. Leg-tracking and automated behavioural classification in *Drosophila*. *Nat. Commun.* 4:1910

72. Berman GJ, Choi DM, Bialek W, Shaevitz JW. 2014. Mapping the stereotyped behaviour of freely moving fruit flies. *J. R. Soc. Interface* 11:20140672

73. Calhoun AJ, Pillow JW, Murthy M. 2019. Unsupervised identification of the internal states that shape natural behavior. *Nat. Neurosci.* 22:2040–49

74. Croft DP, James R, Krause J. 2008. *Exploring Animal Social Networks*. Princeton, NJ: Princeton Univ. Press

75. Krause J, James R, Franks DW, Croft DP. 2015. *Animal Social Networks*. Oxford, UK: Oxford Univ. Press

76. Lukeman R, Li YX, Edelstein-Keshet L. 2010. Inferring individual rules from collective behavior. *PNAS* 107:12576–80

77. Katz Y, Tunstrøm K, Ioannou CC, Huepe C, Couzin ID. 2011. Inferring the structure and dynamics of interactions in schooling fish. *PNAS* 108:18720–25

78. Herbert-Read JE, Perna A, Mann RP, Schaerf TM, Sumpter DJT, Ward AJW. 2011. Inferring the rules of interaction of shoaling fish. *PNAS* 108:18726–31

79. Herbert-Read JE, Krause S, Morrell LJ, Schaerf TM, Krause J, Ward AJW. 2013. The role of individuality in collective group movement. *Proc. R. Soc. B* 280:20122564

80. Jolles JW, Boogert NJ, Sridhar VH, Couzin ID, Manica A. 2017. Consistent individual differences drive collective behavior and group functioning of schooling fish. *Curr. Biol.* 27:2862–68.e7

81. Jolles JW, King AJ, Killen SS. 2020. The role of individual heterogeneity in collective animal behaviour. *Trends Ecol. Evol.* 35:278–91

82. Nagy M, Ákos Z, Biro D, Vicsek T. 2010. Hierarchical group dynamics in pigeon flocks. *Nature* 464:890–93

83. Shannon C, Weaver W. 1998. *The Mathematical Theory of Information*. Urbana: Univ. Ill. Press

84. Schreiber T. 2000. Measuring information transfer. *Phys. Rev. Lett.* 85:461–64

85. Lord WM, Sun J, Ouellette NT, Bollt EM. 2016. Inference of causal information flow in collective animal behavior. *IEEE Trans. Mol. Biol. Multi-Scale Commun.* 2:107–16

86. Sun J, Bollt EM. 2014. Causation entropy identifies indirect influences, dominance of neighbors and anticipatory couplings. *Phys. D* 267:49–57

87. Pilkiewicz KR, Lemasson BH, Rowland MA, Hein A, Sun J, et al. 2020. Decoding collective communications using information theory tools. *J. R. Soc. Interface* 17:20190563

88. Leonard NE. 2014. Multi-agent system dynamics: bifurcation and behavior of animal groups. *Annu. Rev. Control* 38:171–83

89. Eyjolfsdottir E, Branson K, Yue Y, Perona P. 2016. Learning recurrent representations for hierarchical behavior modeling. arXiv:1611.00094 [cs.AI]

90. Shahrokhi S, Becker AT. 2016. Object manipulation and position control using a swarm with global inputs. In *2016 IEEE International Conference on Automation Science and Engineering*, pp. 561–66. Piscataway, NJ: IEEE

91. Mguni D, Jennings J, de Cote EM. 2018. Decentralised learning in systems with many, many strategic agents. arXiv:1803.05028 [cs.MA]

92. Gebhardt GHW, Daun K, Schnaubelt M, Neumann G. 2018. Learning robust policies for object manipulation with robot swarms. In *2018 IEEE International Conference on Robotics and Automation*, pp. 7688–95. Piscataway, NJ: IEEE

93. Zaheer M, Kottur S, Ravanbakhsh S, Poczos B, Salakhutdinov RR, Smola AJ. 2017. Deep sets. In *Advances in Neural Information Processing Systems 30*, ed. I Guyon, UV Luxburg, S Bengio, H Wallach, R Fergus, et al., pp. 3391–401. Red Hook, NY: Curran

94. Hüttenrauch M, Šošić A, Neumann G. 2019. Deep reinforcement learning for swarm systems. arXiv:1807.06613 [cs.MA]

95. Gebhardt GHW. 2019. *Using mean embeddings for state estimation and reinforcement learning*. PhD Thesis, Tech. Univ. Darmstadt, Darmstadt, Ger.

96. Aberdeen D. 2003. *A (revised) survey of approximate methods for solving partially observable Markov decision processes*. Tech. Rep., Natl. Inf. Commun. Technol. Aust., Canberra

97. Wierstra D, Foerster A, Peters J, Schmidhuber J. 2007. Solving deep memory POMDPs with recurrent policy gradients. In *Artificial Neural Networks: ICANN 2007*, ed. JM de Sá, LA Alexandre, W Duch, D Mandic, pp. 697–706. Berlin: Springer

98. Schmidhuber J. 2015. Deep learning in neural networks: an overview. *Neural Netw.* 61:85–117

99. Tobin J, Fong R, Ray A, Schneider J, Zaremba W, Abbeel P. 2017. Domain randomization for transferring deep neural networks from simulation to the real world. arXiv:1703.06907 [cs.RO]

100. Fu J, Levine S, Abbeel P. 2016. One-shot learning of manipulation skills with online dynamics adaptation and neural network priors. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4019–26. Piscataway, NJ: IEEE

101. Isele D, Rostami M, Eaton E. 2016. Using task features for zero-shot knowledge transfer in lifelong learning. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, pp. 1620–26. New York: AAAI Press

102. Finn C, Abbeel P, Levine S. 2017. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning*, pp. 1126–35. Proc. Mach. Learn. Res. Vol. 70. N.p.: PMLR

103. Gilpin LH, Bau D, Yuan BZ, Bajwa A, Specter M, Kagal L. 2018. Explaining explanations: an overview of interpretability of machine learning. In *2018 IEEE 5th International Conference on Data Science and Advanced Analytics*, pp. 80–89. Piscataway, NJ: IEEE

104. Réale D, Reader SM, Sol D, McDougall PT, Dingemanse NJ. 2007. Integrating animal temperament within ecology and evolution. *Biol. Rev.* 82:291–318

105. Freund J, Brandmaier AM, Lewejohann L, Kirste I, Kritzler M, et al. 2013. Emergence of individuality in genetically identical mice. *Science* 340:756–59

106. Bierbach D, Laskowski KL, Wolf M. 2017. Behavioural individuality in clonal fish arises despite near-identical rearing conditions. *Nat. Commun.* 8:15361

107. Deisenroth M, Rasmussen CE. 2011. PILCO: a model-based and data-efficient approach to policy search. In *Proceedings of the 28th International Conference on Machine Learning*, pp. 465–72. Madison, WI: Omnipress

108. Gal Y, McAllister R, Rasmussen CE. 2016. *Improving PILCO with Bayesian neural network dynamics models*. Paper presented at the Data-Efficient Machine Learning Workshop, 33rd International Conference on Machine Learning, New York, June 19–24

109. Ha D, Schmidhuber J. 2018. World models. arXiv:1803.10122 [cs.LG]

110. Oudeyer PY, Kaplan F, Hafner VV. 2007. Intrinsic motivation systems for autonomous mental development. *IEEE Trans. Evol. Comput.* 11:265–86

111. Pathak D, Agrawal P, Efros AA, Darrell T. 2017. Curiosity-driven exploration by self-supervised prediction. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 488–89. Piscataway, NJ: IEEE

112. Schmidhuber J. 2006. Developmental robotics, optimal artificial curiosity, creativity, music, and the fine arts. *Connect. Sci.* 18:173–87

113. Schmidhuber J. 2007. Simple algorithmic principles of discovery, subjective beauty, selective attention, curiosity & creativity. In *Discovery Science: 10th International Conference, DS 2007*, ed. V Corruble, M Takeda, E Suzuki, pp. 26–38. Berlin: Springer

114. Schmidhuber J. 2010. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Trans. Auton. Ment. Dev.* 2:230–47

115. Couzin ID, Krause J, Franks NR, Levin SA. 2005. Effective leadership and decision-making in animal groups on the move. *Nature* 433:513–16

116. Ioannou C, Singh M, Couzin I. 2015. Potential leaders trade off goal-oriented and socially oriented behavior in mobile animal groups. *Am. Nat.* 186:284–93

117. Frohnwieser A, Murray JC, Pike TW, Wilkinson A. 2016. Using robots to understand animal cognition: robots in animal cognition. *J. Exp. Anal. Behav.* 105:14–22

118. Goodale E, Beauchamp G, Magrath RD, Nieh JC, Ruxton GD. 2010. Interspecific information transfer influences animal community structure. *Trends Ecol. Evol.* 25:354–361

119. Bieker K, Peitz S, Brunton SL, Kutz JN, Dellnitz M. 2019. Deep model predictive control with online learning for complex physical systems. arXiv:1905.10094 [cs.LG]

120. Wülfing JM, Kumar SS, Boedecker J, Riedmiller M, Egert U. 2019. Adaptive long-term control of biological neural networks with deep reinforcement learning. *Neurocomputing* 342:66–74

121. Tessadori J, Chiappalone M. 2015. Closed-loop neuro-robotic experiments to test computational properties of neuronal networks. *J. Vis. Exp.* 97:42341