Title
Robots Can Train Humans Using Principles of Operant Conditioning Through Visual Reinforcement Tools

Journal
International Journal of Comparative Psychology, 31(0)

ISSN
0889-3675

Authors
Kieson, Emily
Maske, Harshal
Abramson, Charles Ira
et al.

Publication Date
2018

DOI
10.46867/ijcp.2018.31.04.01

INTERNATIONAL JOURNAL OF COMPARATIVE PSYCHOLOGY

2018, 31
Heather M. Hill
Editor
Peer-reviewed

Robots Can Train Humans Using Principles of Operant Conditioning Through Visual Reinforcement Tools

Emily Kieson[1], Harshal Maske[2], Charles I. Abramson[1], Girish Chowdhary[2], and Christopher Crick[3]

[1]Department of Psychology, Oklahoma State University, U.S.A.
[2]Department of Agricultural and Biological Engineering, University of Illinois, U.S.A.
[3]Computer Science Department, Oklahoma State University, U.S.A.

Researchers have established new techniques to study human-robot interactions based on current knowledge in interspecies communication and comparative psychology. Studies on animal acceptance of robot conspecifics in complex social environments has led to the development of robots that adapt to animal and human behaviors. Using a robot with adaptable algorithms developed by the authors, the researchers hypothesized that, by using familiar visual rewards as positive reinforcement, robots could use operant conditioning principles to teach humans a basic task. The robot in this study independently determined optimal control of construction equipment by capturing the motions from an expert operator. The robot then attempted to teach those same skills to novice operators using familiar, yet simple, visual reinforcement tools. In this study, participants were asked to manipulate a model excavator using feedback from the guidance system on a nearby computer screen. Participants were assigned randomly to one of 3 groups: simple visual reinforcement, complex guidance, and no visual feedback (blank screen). To measure learning, participants returned a day later to repeat the task without the guidance. The group using simple feedback resulted in cycle times that were closer to the expert times than both the complex or control groups and had significantly different end times ($p < 0.05$) than either group. This result supports the hypothesis that, similar to what has been found in vertebrates and invertebrates, robots can shape behaviors of humans using visual positive reinforcement.

Keywords: robots, learning, operant conditioning

Comparative psychology has a rich history of using apparatuses and standardized methods to study behavioral responses in animals and comparing these responses to how humans relate to each other and the world (Varnon, Lang, & Abramson, 2018). Prior to the introduction of robotics, comparative psychology and anthrozoology provided a solid foundation for studying how humans interact with other species. With the advent of robotics, this new field created opportunities for integrating comparative psychology and interspecies communications into the ever-evolving world of human-robot interactions (Abramson, 2018).

The research in human-animal interactions serves as a model for human-robot interactions through behavioral feedback and applied learning theory that occurs during inter-species communication (Korondi, Korcsok, Kovács, & Niitsuma, 2015). For example, researchers have used information on human-dog relationships as a basis for developing more socially-acceptable robots that mimic human social patterns (Konok, Korcsok, Miklósi, & Gácsi, 2018). Additionally, robots have been designed for animal ethological research in which the robot was designed and accepted as a conspecific (Gribovskiy, Halloy, Deneubourg, & Mondada, 2018), proving that animals can incorporate robots into their daily lives if the robot is equipped with the proper tools for behavioral interactions. For these same principles to be applied to human-robot research, researchers need to integrate socially acceptable interfaces so that the robots conform to expected standards of interaction that meet the norms of human-animal or human-human interactions. Interactive behavioral models of interspecies communication between humans and other animals are therefore critical in understanding the human-robot interaction.

Please send correspondence to Charles I. Abramson, Oklahoma State University, U.S.A. (Email: charles.abramson@okstate.edu)
https://doi.org/10.46867/ijcp.2018.31.04.01

To facilitate the integration of robots into society, scientists have developed the field of ethorobotics, which uses behavioral models to create mathematical principles for socially interactive robots (Korondi et al., 2015). This field provides a unique approach to studying behavior in both humans and animals because their interactions can be standardized and adapted based on human and animal feedback (Krause, Winfield, & Deneubourg, 2011). These adaptive functions provide insight into how robots can learn and provide individualized feedback based on changes in human interactions in a learning environment.

In addition to having socially acceptable interactions, cooperative robots also need to be able to adapt to human needs in ways that are comparable to complex social structures indicative of many species. Complex social structures in animal species often result in specialization of individuals, cooperative behavior, and communication skills necessary to convey useful information. Although socially acceptable interactions are important, robots also need to be able to adapt to human interactions. Principles of teamwork in invertebrates and other animals also apply to humans, especially concerning cooperative social structures (Anderson & Franks, 2003). If robots are to be incorporated into a social system, scientists need to ensure that robots conform to the same social expectations and interpersonal dependence as other members of the team. As robots become more complex and adaptive, there is greater potential for them to adapt and carry out the required communication necessary to convey and receive information from humans and create new response patterns that match the needs of the environment (Gigliotta, 2018). If robots are to fill a new niche in human culture, they, too, need to be adaptive and find unique ways of communicating with humans that fit with evolved cooperative behavioral strategies already in place (Gigliotta, 2018). This kind of automated adaptability is essential for continuing to provide immediate behavioral and environmental feedback needed to train or shape behaviors in the animal or human subject.

The required social complexity, socially acceptable interactions, and adaptability of robotic partners are essential for cooperative robots to adhere to existing behavioral and learning theories. Past research in comparative psychology of insects has provided engineers with not just physiological inspiration for mechanical design, but also opportunities to mimic basic behavioral and learning models based on invertebrate systems (Webb, 2017). Insect physiology and basic learning feedback mechanisms found in comparative psychology have given rise to small adaptive robots capable of maneuvering and adapting under very specific environments (Webb, 2017). Such specially-programmed robots are also capable of teamwork and assessing and adapting behaviors in order to work together in a manner consistent with cooperation in biological species (Anderson & Franks, 2003). Robots have also been designed to mimic the behavioral and neurological processing of insects in order to better understand the brain patterns and predict behaviors in insects (Szczecinski, Getsy, Martin, Ritzmann, & Quinn, 2017), which gives rise to robots that can mimic animal behaviors and adapt in similar (basic) environments. These types of bio-robots have also provided researchers with tools to create predictive models of movement and behavior based on comparisons between biological organisms and their robotic counterparts (Ayali et al., 2015).

The robotic adaptations used in animal interactions are also applicable to humans. The purpose of this study is to help define how principles of learning theory can be applied in human-robot interactions through existing reward mechanisms for application in industrial settings. For this study, the researchers focused on an adaptable robotic interface capable of interacting with human participants in a way that mimicked common social and environmental interactions to determine the impact of cooperative robots on human learning and behavioral shaping. To mimic visual and culturally ingrained reinforcement, the researchers chose easily recognizable red and green colors to indicate correct (green) or incorrect (red). These colors also come with subtle cultural associations of good and bad (Gilbert, Fridlund, & Lucchina, 2016; Gong, Wang, Hai, & Shao, 2017), which helps to create the emotional associations needed to maintain reinforcement values. The rationale for using colors was to manipulate existing familiar visual reinforcement options to shape the behaviors of the participants. Because red and green are already familiar and have existing associations within the cultural context, these were easily utilized to provide the necessary positive reinforcement.

The hypotheses were that robots could teach humans a complex task with basic operant conditioning by using green color indicators on a screen as a visual positive reinforcement indicator when the task had been performed correctly and that this basic conditioning would result in greater learning than the more complex visual guidance interfaces used on the same screen.

Method

Participants

A total of 113 participants volunteered for the study or were recruited through the university research participation system (SONA). The average age of participants was 23.83 years (SD = 5.50 years) with approximately 72% males and 27% females, and 26.5% Asians, 33.6% Caucasian, and the remainder of varying self-reported ethnic descent. All volunteers were active students, faculty, or visitors in the psychology and engineering departments at Oklahoma State University in Stillwater, Oklahoma.

Measures

Participant actions were evaluated based on the number of times they hit the truck with the bucket of the excavator (*truck hits*), how many single motions it took for the participant to maneuver the excavator into the correct position (*mean actions*), and how many times they needed to correct their movements (*errors*). Each of these were counted based on video analysis of the participants and compared between groups, signifying how closely the participant behavior matched that of the expert with regards to joystick movement. In addition to the physical motions of the excavator, the timing of each trial was calculated for each individual, and the means were then calculated within each group for each trial. The number of truck hits, mean actions, errors, and the trial times were calculated as continuous variables and processed through SPSS to determine significance based on appropriate statistical analyses.

Procedure

Upon arrival, participants were randomly assigned to one of three groups to test the hypothesis that a basic graphical user interface (GUI) could be used for robots to teach human learners. Group 1 consisted of 34 participants, Group 2, 41 participants, and Group 3, 38 participants. Participants were randomly sorted into groups: Group 1 using the GUI with circles, Group 2 using the GUI with speed bars, and Group 3 not using a GUI (control).

Once the participants finished the initial introduction, a research assistant introduced them to the GUI on the computer screen and explained what the interface meant. All participants sat in a standard office chair and viewed a guidance system on a standard 15-in. computer screen set on a box slightly to their left as to not block the view of the remote-controlled excavator located directly in front of them on the floor. Group 1 with the colored circles (see Figure 1) experienced changes from red to green circles for each hand if the actions were congruent with expected or optimal actions. Group 2 had speed bars (Figure 2) that integrated the color reinforcements of Group 1 with direction arrows giving more complex visual cues as to what the participants should do with regards to the joysticks. Group 3 was the control group and had blank screens.

Each participant was given a minimum of three trials to attempt to scoop sand from the tub and deposit into the truck using the controls and the assistance of the selected training GUI. Participants were timed and videoed, and actions and times were then compared to ideals set in place by “experts” who had extensive experience and familiarity with the equipment. Participants then returned within three days to repeat the trial, but this time without any help of a GUI. Measurments for time, errors and actions were taken on both the initial and return visits.

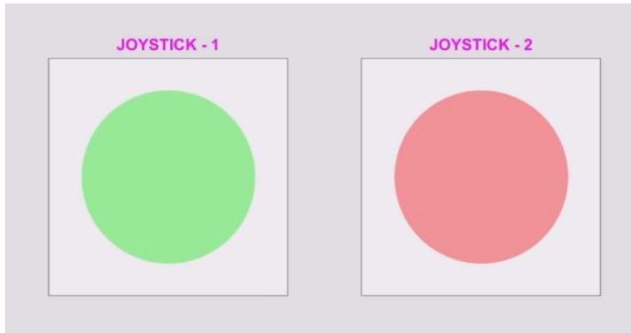

Figure 1. **Guidance User Interface (GUI) for Group 1 – Circles.**

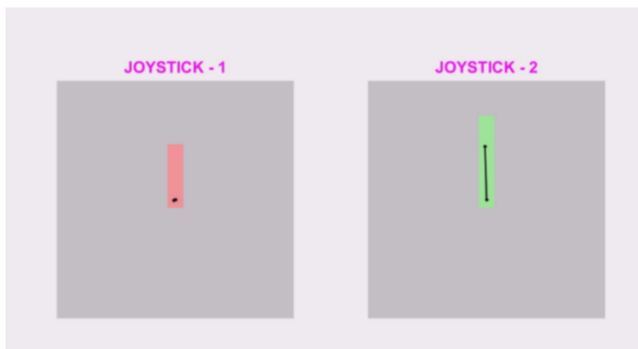

Figure 2. **Guidance User Interface (GUI) for Group 2 – Speed Bars.**

A flow diagram of the visual interface is shown in Figure 3. For this study, the researchers used a co-robot that utilized real-time feedback through the joystick operated by the participant. The position and movement of the excavator was collected through position markers and digital cameras in the room, which were then relayed to the co-robot along with the joystick movements of the participant. The co-robot then provided visual feedback through the GUI based on real-time data of the excavator position and joystick movements. The co-robot was programmed with previous input from experienced operators and, based on these known "optimal" parameters for position and speed, the co-robot adjusted the visual feedback to help guide the participant to adjust his or her joystick movements to best match those of the experienced operators. The instructional policy algorithm created for this part of the study was based on research in artificial intelligence, reinforcement learning, and learning from demonstration (Maske, Kieson, Chowdhary, & Abramson, 2016).

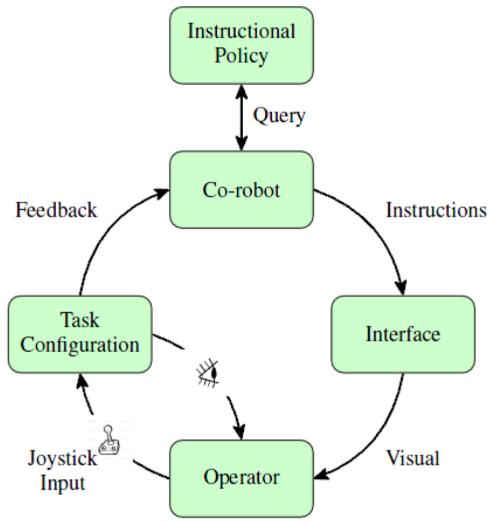

Figure 3. **Instruction Interface Flow Diagram.**

The remote-controlled excavator was a fully hydraulic 1/14th scale model (Figure 4) of 345D CAT Excavator similar to the model made by Wedico and available through Amazon.com. The model lacked joint-angle encoders and internal proprioception, hence all the experiments were performed inside a motion capture facility to ensure individual joint movements are captured in real-time and compared with optimal trajectories and actions. The motion capture facility included small physical markers on each joint of the excavator as well as interaction points for the cameras to locate the position of the sand bin and the truck. Ceiling-mounted cameras tracked the motion of the pins and relayed it back to the computer so that the computer could calculate positions, angle, and trajectory during each motion in relation to how the user manipulated the joystick.

Figure 4. **Model Excavator and Truck.**

Results

Participants were assigned to one of three groups, each with a different GUI and are categorized as follows: Group 1 – GUI Circles, Group 2 – GUI Speed Bars, and Group 3 – No GUI.

Cycle Times

Each participant was asked to complete the task at least three times and each cycle was timed. Means were calculated for each cycle/trial for each group for both original and retest cycle times. Paired *t* tests were then conducted to compare the times of Cycle 1 to Cycle 3, Cycle 3 to Retest 1, and Retest 1 to Retest 3. An ANOVA was used to compare the timed results of Retest 3 (Table 1) and a one-sample *t* test was used to compare Retest 3 times with the expert time (24.9 s).

Analysis of the distribution of data showed extreme outliers for Group 1 and Group 3 so data were selected to only include retest times that were less than 65 s. Two cases were removed from Group 1 (65 s, 140 s) and one case was removed from Group 3 (85 s).

The results of the paired-samples *t* test for cycle times suggested that there were no significant differences between times for Cycle 1 and Cycle 3 or between Cycle 3 and Retest 1. Paired-samples *t* tests for retest times resulted in significant changes between Retest 1 and Retest 3 times for both Group 1 ($M = 18.62$, $SD = 22.16$), $t(28) = 4.53$, $p < 0.01$ $d = 1.04$, and Group 3 ($M = 10.76$, $SD = 22.29$), $t(28) = 2.60$, $p < 0.05$, $p = .015$, $d = 0.54$, suggesting that Group 1 and Group 3 had more significant changes in retest cycle times than did Group 2.

An ANOVA on Retest 3 times showed significant differences between at least two of the groups, $F(2, 81) = 6.41$, $p < 0.01$. A Tukey post-hoc was performed and resulted in significant differences between Retest 3 times for Group 1 and Group 2, $p < 0.01$, and between Group 1 and Group 3, $p = 0.008$, suggesting that those who used the circle GUI (Group 1) demonstrated significant differences in their Retest 3 times when compared to the other two groups.

Figure 5 shows the mean times for each group with corresponding error bars at Cycle Time 1, Cycle 3, Retest 1, and Retest 3. The lines joining the means facilitate differentiation between data points and do not refer to any regression analysis.

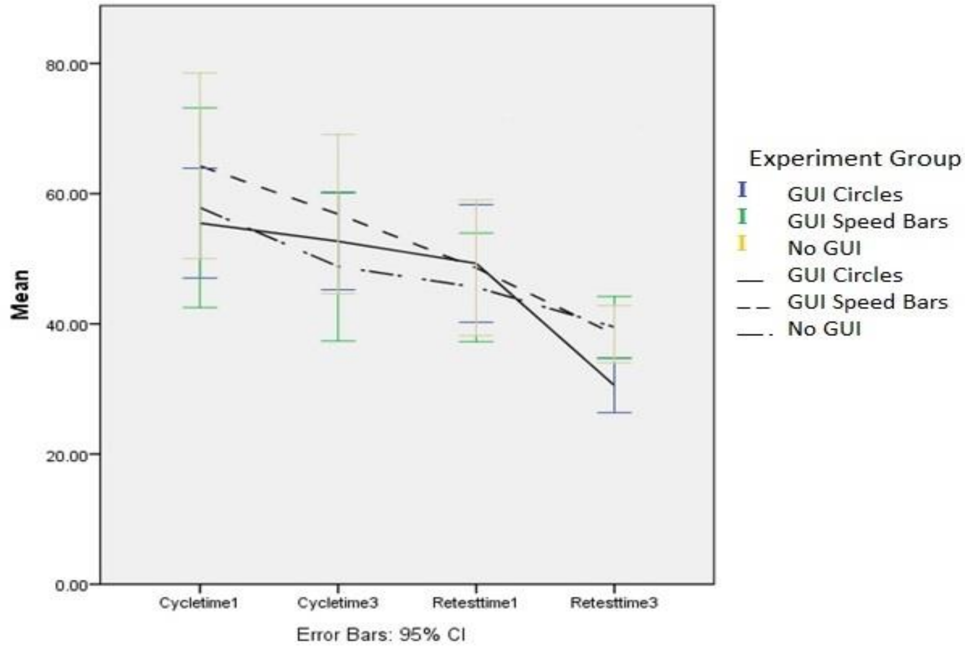

Figure 5. **Mean Cycle Times and Corresponding Errors for Experimental Groups.**

Based on the results, visual interfaces seem to provide some improvement over the control with regards to errors and, although there were significant changes in cycle times for Group 1 (GUI Circles) and Group 3 (No GUI) over the duration of the cycles, all groups remained significantly different than the expert at the end of the retests. Group 1, however, showed significantly different times at the end than the other two and came closest to the expert time of 24.9 s.

Discussion

Overall, the results of this experiment demonstrated that a simple visual reinforcement worked just as effectively as more complex demonstration systems. The independent variables of simple visual reinforcement versus complex demonstration interfaces were chosen to determine if basic visual reinforcement would be more effective than a more complex guidance system. The simple visual reinforcer interface relies on basic principles of operant conditioning, in which the user is guided to repeat behaviors for which he or she has received positive reinforcement. In this study, a green color signified a correct movement of the joystick and a red color signified a wrong movement. The more complex guidance system introduced arrows and directions in addition to color changes to assist the user in performing the task as closely to the expert movements as possible. Because some simulators and video games use more complex systems to teach a task (Luria & Vogel, 2011), it was important to test whether a simple interface could achieve the same result. Based on the end times for the retest, participants using the simple color interface had retest times closer to that of the experts and significantly different than the other groups. This suggests that the basic interface using socially acceptable feedback worked effectively as a reinforcer when training humans.

The simplicity of the interface may be a key component of the learning interaction. In a mixed-method study of video-based classrooms in online learning environments, researchers found that students performed better with basic visual interfaces compared to complex alternatives. The students reported that the elaborate videos were distracting from the lesson due to unnecessary visual elements. The simple videos, on the other hand, were more focused on material rather than elaborate graphics (Guo, Kim, & Rubin, 2014). The use of simple videos in an online learning environment futher supports the findings that a basic graphic interface outperforms a more complex version for both user focus as well as reinforcement learning.

Our findings indicate that the simple visual reinforcement interface outperformed the other interfaces and controls, which suggests that, by understanding learning theory, we can develop co-robots that utilize basic reinforcement strategies to shape the behaviors of humans. Considering that similar studies have been done with insects, in which robots interacted with insects through reinforcement learning and adapted strategies based on insect behavioral feedback (Son, Choi, & Ahn, 2014), this principle could potentially apply to many areas of robot-animal interactions.

Although the study supports the use of reinforcement learning in robot-human interactions, future studies should incorporate the use of additional control groups as well as considerations for personality traits and their correlations to potential study outcomes. The study used only volunteers from departments within Oklahoma State University and would have benefited from the use of individuals with experience in construction. In addition, the use of standardized personality questionnaires may help to clarify differences (or correlations) between university volunteers and the target application market of construction workers in order to provide a more thorough understanding of how the outcomes could pertain to different populations. There are plans to study the use of this system in construction training simulators and life-size equipment to collect data more directly applicable to the intended audience, and those results will be compared to the current findings.

The development of robots as socially interactive machines gives researchers more insight into how we can compare the development of robotic behaviors and artificial intelligence with that of humans. This new insight both compares and contrasts robots and humanity (Cockshott & Renaud, 2016), especially because robots can learn to model behavior and therefore test the successful implementation of different teaching styles (Krause et al., 2011). By dissecting the nature of human behavior and comparing it to that of robots in a social and teaching context, we begin to evaluate the processes that define biological intelligence from artificial processes (Cockshott & Renaud, 2016). This is especially important as scientists and engineers continue to develop robots for both industry and personal use, further defining their place in society through scientific evaluation (Cockshott & Renaud, 2016).

As the roles of robots change, it is important to direct research efforts to determine how to best facilitate the transition of robots from learners into teachers. Industrial robots used as part of a workforce are usually handled by individuals with a knowledge of robot management and maintenance. Cooperative social robots, however, like the ones used to train humans, are typically used as a tool for training or support. These cooperative robots typically interact with humans who are less skilled in robotics, which requires robots to interact with humans in ways that align with normal human interactions and conform to social norms (Korondi et al., 2015).

The co-robots in this study easily communicated correct and incorrect motions to participants using only basic color changes, suggesting that the use of basic reinforcement tools may work well for cooperative robots. With regards to the future of robots in society, our results suggest that robots can more readily train novice humans to perform any variety of tasks simply by manipulating existing reinforcers. Robots and, more importantly, artificial intelligence continue to grow in importance as their development affects how we work and live. By understanding how robots interact with animals and vice versa, researchers are now beginning to explore how different species may incorporate robots into daily life. Studies in robot-animal and robot-human interactions give rise to new and exciting opportunities in comparative psychology to increase our understanding of interspecies communications and may provide a critical foundation in the development of adaptable robots and artificial intelligence.

References

Abramson, C. I. (2018). Let us bring comparative psychology back. *International Journal of Comparative Psychology*, *31*. Retrieved from https://escholarship.org/uc/item/81j662cd.

Anderson, C., & Franks, N. R. (2003). Teamwork in animals, robots, and humans. *Advances in the Study of Behavior, 33*, 1–48. http://doi.org/10.1016/S0065-3454(03)33001-3

Ayali, A., Borgmann, A., Büschges, A., Couzin-Fuchs, E., Daun-Gruhn, S., & Holmes, P. (2015). The comparative investigation of the stick insect and cockroach models in the study of insect locomotion. *Current Opinion in Insect Science*, *12*, 1–10. http://doi.org/10.1016/j.cois.2015.07.004

Cockshott, P., & Renaud, K. (2016). Humans, robots and values. *Technology in Society*, *45*, 19–28. http://doi.org/10.1016/j.techsoc.2016.01.002

Gigliotta, O. (2018). Equal but different: Task allocation in homogeneous communicating robots. *Neurocomputing, 272*, 3–9. http://doi.org/10.1016/j.neucom.2017.05.093

Gilbert, A. N., Fridlund, A. J., & Lucchina, L. A. (2016). The color of emotion: A metric for implicit color associations. *Food Quality and Preference, 52*, 203–210. http://doi.org/10.1016/j.foodqual.2016.04.007

Gong, R., Wang, Q., Hai, Y., & Shao, X. (2017). Investigation on factors to influence color emotion and color preference responses. Optik - *International Journal for Light and Electron Optics, 136*, 71–78. http://doi.org/10.1016/j.ijleo.2017.02.026

Guo, P.J., Kim, J., Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. *Proceedings of the first ACM Conference on Learning @ Scale Conference,* 41 – 50. https://doi.org/10.1145/2556325.2566239

Gribovskiy, A., Halloy, J., Deneubourg, J. L., & Mondada, F. (2018). Designing a socially integrated mobile robot for ethological research. *Robotics and Autonomous Systems, 103*, 42–55. http://doi.org/10.1016/j.robot.2018.02.003

Konok, V., Korcsok, B., Miklósi, Á., & Gácsi, M. (2018). Should we love robots? – The most liked qualities of companion dogs and how they can be implemented in social robots. *Computers in Human Behavior, 80*, 132–142. http://doi.org/10.1016/j.chb.2017.11.002

Korondi, P., Korcsok, B., Kovács, S., & Niitsuma, M. (2015). Etho-robotics: What kind of behaviour can we learn from the animals? *IFAC-PapersOnLine, 48*, 244–255. http://doi.org/10.1016/j.ifacol.2015.12.041

Krause, J., Winfield, A. F. T., & Deneubourg, J. L. (2011). Interactive robots in experimental biology. *Trends in Ecology and Evolution, 26*, 369–375. http://doi.org/10.1016/j.tree.2011.03.015

Luria, R., & Vogel, E. K. (2011). Shape and color conjunction stimuli are represented as bound objects in visual working memory. *Neuropsychologia, 49*, 1632–1639. http://doi.org/10.1016/j.neuropsychologia.2010.11.031

Maske, H., Kieson E., Chowdhary, G., & Abramson, C. I. (2016). Can co-robots learn to teach? arXiv.org.arXiv:1611.07490.

Son, J. H., Choi, Y. C., & Ahn, H. S. (2014). Bio-insect and artificial robot interaction using cooperative reinforcement learning. *Applied Soft Computing Journal, 25*, 322–335. http://doi.org/10.1016/j.asoc.2014.09.002

Szczecinski, N. S., Getsy, A. P., Martin, J. P., Ritzmann, R. E., & Quinn, R. D. (2017). Mantisbot is a robotic model of visually guided motion in the praying mantis. *Arthropod structure and development, 46*, 736–751. http://doi.org/10.1016/j.asd.2017.03.001

Varnon, C. A, Lang, H., & Abramson, C. I. (2018). Automated research in comparative psychology: Limitations and new directions. *International Journal of Comparative Psychology, 31*. Retrieved from https://escholarship.org/uc/item/8gf4d5tx

Webb, B. (2017). From insects to robots. *Arthropod structure and development, 46*, 687–688. http://doi.org/10.1016/j.asd.2017.08.002

Financial conflict of interest: This study was possible through a grant by the National Science Foundation and the National Robotics Initiative (1527828).

Conflict of interest: No stated conflicts.

Submitted: *June 19th, 2018*

Resubmitted: *July 18th, 2018*

Accepted: *September 14th, 2018*