Neuroscience Research xxx (2019) xxx–xxx

Contents lists available at ScienceDirect

Neuroscience Research

journal homepage: www.elsevier.com/locate/neures

Short Communication

Reinforcement schedules differentially affect learning in neuronal operant conditioning in rats

Kichan Song[a], Susumu Takahashi[b], Yoshio Sakurai[a,*]

[a] Laboratory of Neural Information, Graduate School of Brain Science, Doshisha University, Kyotanabe, 610-0394, Kyoto, Japan
[b] Laboratory of Cognitive and Behavioral Neuroscience, Graduate School of Brain Science, Doshisha University, Kyotanabe, 610-0394, Kyoto, Japan

ARTICLE INFO

Article history:
Received 11 March 2019
Received in revised form 1 April 2019
Accepted 12 April 2019
Available online xxx

Keywords:
Operant conditioning
Neuronal activity
Reinforcement schedule
Brain-machine interface
motor cortex
Rat

ABSTRACT

Operant conditioning of neuronal activity is a core process for better operation of brain-machine interfaces. However, few studies have investigated the role of reinforcement schedules in neuronal operant conditioning, although they are very effective in behavioral operant conditioning. To test the effect of different reinforcement schedules, the authors trained single-neuron activity in the motor cortex using fixed ratio (FR) and variable ratio (VR) schedules in rats. Neuronal firing rates were enhanced in the FR but not in the VR schedule during conditioning, suggesting that the principles of operant conditioning of neuronal activity are different from those of behavioral responses.

Brain-machine interfaces (BMIs) enable neuroprosthetic control of external devices by brain activity alone (Chapin et al., 1999; Nicolelis, 2011; Lebedev, 2014). Despite the steady progress in the development of BMIs for future clinical use (Collinger et al., 2013; Donoghue et al., 2007), their current accuracy and efficiency are limited and, as such, improvements in some technical factors affecting BMI performance are actively being pursued (Ethier et al., 2012; Lebedev and Nicolelis, 2011). However, previous studies (Andersen et al., 2010; Nicolelis and Lebedev, 2009) have emphasized that improvements in technical factors alone cannot solve all problems that hinder the development of BMIs; therefore, increased knowledge of brain mechanisms is needed (Baranauskas, 2014; Mandonnet and Duffau, 2014; Sakurai, 2014; Velliste et al., 2014). In particular, the plasticity of neuronal networks and their activity is essential (Sakurai, 2014), and accurate device control by BMIs inevitably requires neuronal activity to be volitionally modulated (Sakurai et al., 2014). The core process of such volitional modulation is operant conditioning of neuronal activity (Sakurai and Song, 2016), because operating machines with BMIs is performed to achieve some type of goal, and goal achievement functions as a reward that enhances altered neural activity through reinforcement feedback (Fetz, 2007).

Research investigating neuronal operant conditioning began with Fetz (1969) and is steadily becoming more prolific (Arduin et al., 2013; Engelhard et al., 2013; Hwang et al., 2013; Koralek et al., 2012) given its relation to BMIs (Fetz, 2007). However, there have been few systematic investigations of the methodology for operant conditioning of neuronal activity. In particular, it has not been clearly demonstrated how reinforcement schedules affect learning in neuronal operant conditioning for volitional modulation of neuronal activity, although they are known to have a significant effect on an animal's operant responses in behavioral studies (Reynolds, 1975).

In the present study, we focused on the effects of reinforcement schedules applied to single neurons of the motor cortex during neuronal operant conditioning in rats. We used two basic—yet different—schedules of reinforcement: fixed ratio (FR) and variable ratio (VR). Both schedules have been shown to increase the probability of occurrence of behavioral response in many animal species and humans, with the former causing stable and moderate rates of response, and the latter causing rapid and high rates (Reynolds, 1975). We tested whether these effects of FR and VR schedules of reinforcement on behavioral responses exist in operant conditioning of neuronal activity.

Animals and apparatus: Eight male albino Wistar rats (Shimizu Laboratory Supplies, Kyoto, Japan), weighing 410–470 g were used as subjects. The animals were individually housed in a $25 \times 15 \times 20$ cm cages under a light and dark cycle (lights on at

* Corresponding author.
E-mail address: ysakurai@mail.doshisha.ac.jp (Y. Sakurai).

https://doi.org/10.1016/j.neures.2019.04.003
0168-0102/

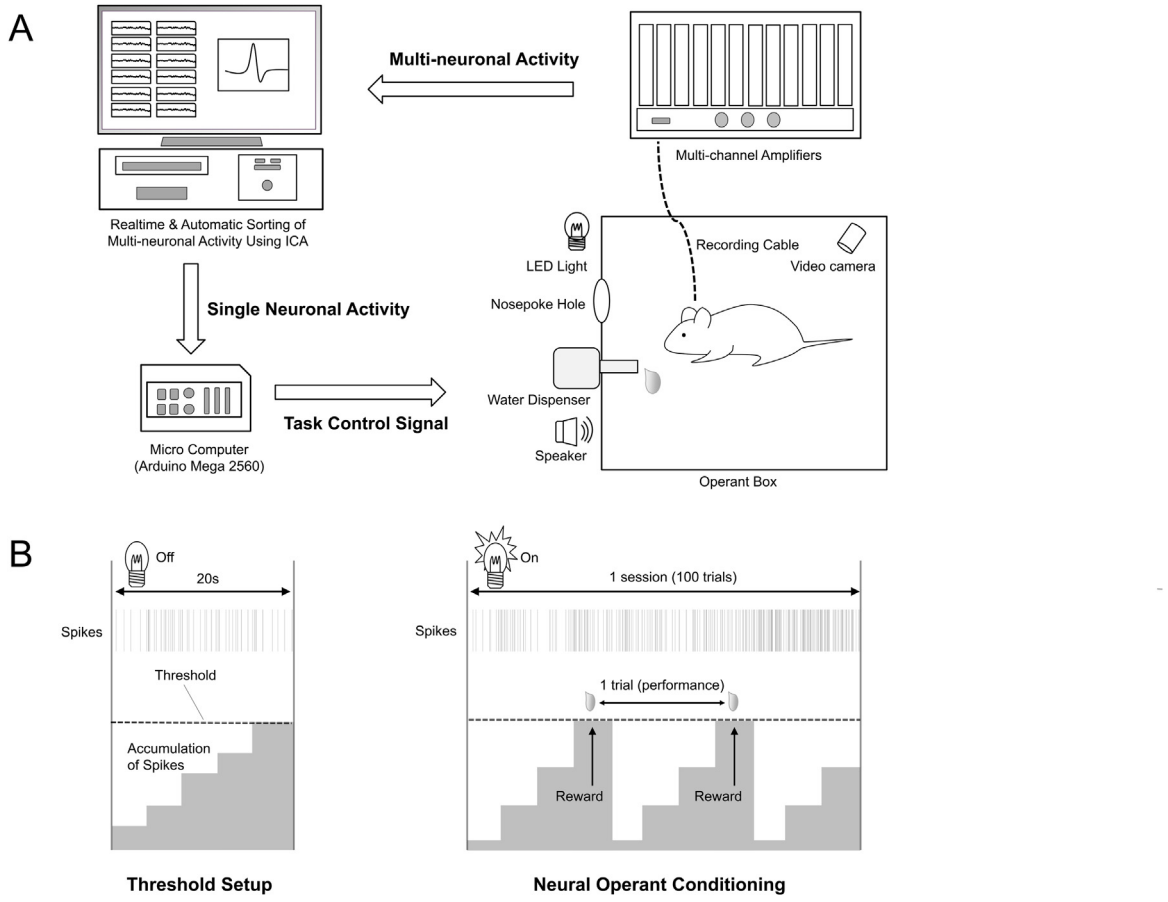

Fig. 1. Schematic diagram of the experimental system and the basic procedures for neuronal operant conditioning.
(A) The system mainly consists of multi-neuronal recording device, a computer for real-time spike sorting with independent component analysis (ICA), and a microcomputer for task controls. The rat is connected to the system through a recording cable and can freely move in the operant box during training sessions. (B) Threshold set-up (left) and flows of neuronal operant conditioning (right). During the threshold set-up, spikes from the target neuron were counted for 20 s.

08:00, lights off at 21:00). The animals' diets were restricted so that their weights were maintained to levels as stable as possible throughout the experiment; however, they had ad libitum access to water except for the behavioral and neuronal operant conditioning periods. Whole training sessions were conducted in an operant box ($22 \times 49 \times 45$ cm) (O'Hara & Co., Tokyo, Japan). A water dispenser controlled by a solenoid valve, a buzzer speaker, a sensor hole, and a light-emitting diode (LED) light were installed on the same side of the operant box, and were controlled by a microcomputer system (Arduino LLC, Italy) (Fig. 1A). As reward, a droplet of water was delivered through a pipe connected to the water dispenser. The volume of the reward was varied (0.04–0.05 ml) depending on the weight of each rat. The buzzer sounded for 1 s when the water emerged from the pipe so that the rats could immediately notice delivery of the reward, regardless of their location or direction in the operant box. The rats' behavior was monitored and recorded using a video camera. In addition, in order to precisely detect rat's bodily movements, a method of three-axis accelerometer (MPU-6050; InvenSense Inc., USA) was used for three of the rats (see Robert et al., 2009 for details). The accelerometer was attached to the connector of the recording cable. Acceleration data was acquired at 10 Hz of sampling rate within a range of ±4 gravitational force (g). The movements of the rats were recorded for 3 s prior to the reward delivery and assessed as the signal vector magnitude (SVM) of the acceleration value (α) from each axis (x, y, z).

$$\mathrm{SVM} = \sqrt{x_a{}^2 + y_a{}^2 + z_a{}^2}$$

All experiments were performed in accordance with guidelines for animal experiments at Doshisha University.

Behavioral operant conditioning: Simple behavioral operant conditioning, in which rats could learn the basic rule of operant conditioning, was used. At first, water was automatically delivered at fixed intervals of 20 s and was accessible by the rats. The rats were then trained to poke their nose into the sensor hole (i.e., nose-poke response) to obtain the reward. Finally, the LED light was turned on and off periodically at 5 min intervals, and the reward was delivered only in response to nose pokes when the LED light was on. Through this operant conditioning procedure, the rats learned that voluntary nose-poke responses during the "LED on" periods led to reward. The operant conditioning protocol was generally completed in a day when the rats performed > 50 nose-poke responses in < 3 min during the LED-on condition. After behavioral training, surgery was performed for electrode implantation.

Electrode implantaion: The electrodes and microdrives were essentially identical to those used in previous studies by the authors (Takahashi and Sakurai, 2005, 2007, 2009a, 2009b; Sakurai and Takahashi, 2013), except that each electrode bundle comprised 6 or 12 tungsten microwires (12.5 μm diameter; California Fine Wire, CA, USA). After completion of behavioral operant conditioning, the electrode bundles were surgically implanted into the rat's primary motor cortex (+ 3.4 mm from bregma, 3.2 mm from mid-line) under anesthesia (isoflurane, approximately 2.5%).

Recording and spike detection: After recovery, multi-neuronal activity from the rats was recorded extracellularly, and spiking activity from individual neurons was detected using the origi-

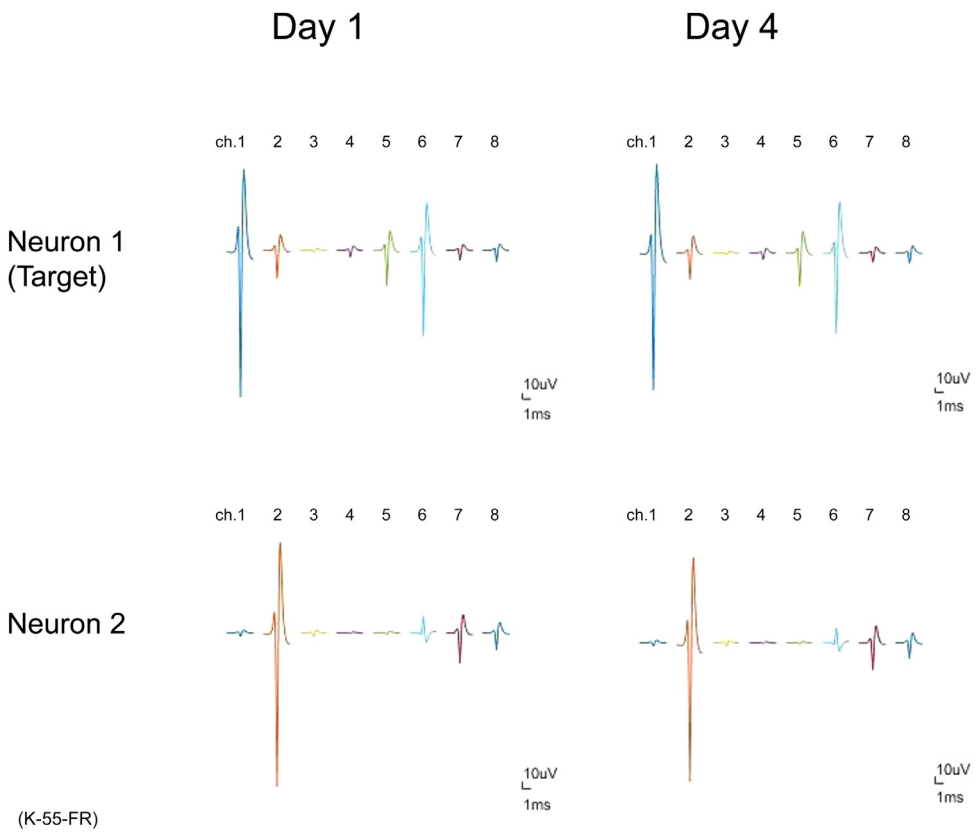

Fig. 2. An example of the activity of separated two single neurons of the motor cortex during the first day (Day 1) and the last day (Day 4) of conditioning.
The averaged spike waveforms on each channel (microwires 1–8 in this example) of the electrode bundle are shown. Each row of 8 waveforms is judged to represent a single neuron by the spike-sorting system. All spikes in each row including very small ones contributed to the calculation of spike-sorting (see Takahashi and Sakurai, 2005; Sakurai and Takahashi, 2013 for detail). The single neuron represented by the first row was the target neuron used for conditioning.

nal system for real-time and automatic sorting of multi-neuronal activity using independent component analysis (ICA) (Takahashi and Sakurai, 2005, 2013). Stable single-neuron activity for neural operant conditioning was identified by lowering the tips of the electrodes approximately 90 μm per day. Three criteria were set for recruiting suitable neurons (target neurons) for conditioning: the amplitude of each spike from a neuron was constant for at least 2 successive days; the mean firing rate was 3–8 Hz; and, finally, the signal-to-noise ratio was > 3. Only neurons that fulfilled all criteria were selected as targets for neuronal operant conditioning (Fig. 2).

Neuronal operant conditioning: Spikes from each of the selected target neurons were counted as the rats moved freely, and a threshold was set for each target neuron for reward delivery. The recording time to obtain spike data to determine thresholds was 20 s and the bin to count spike occurrences was 20 min. The number of spikes counted during the 20 s period were averaged per second (3.0–8.2, mean 5.2) and used as the threshold. During a trial of neuronal operant conditioning, the rats were rewarded whenever the number of spikes from the target neurons reached a "reinforcement threshold" (Fig. 1B). In the FR schedule, the values of reinforcement threshold were fixed and identical with previously determined thresholds. In the VR schedule, the values of reinforcement threshold were not fixed, but randomly varied from trial to trial, ranging from 60, 80, 100, 120, and 140% of the previously determined thresholds so that the rats could not predict reward delivery. The means of varied reinforcement thresholds in each session were always identical to those in the FR schedule (i.e., previously determined thresholds). The total number of spikes for reward in the VR schedule was set to be almost identical with that in the FR schedule for each neuron. Following reward delivery, a 1 s ingestion period was imposed, and the spike counter was reset to zero and the next trial started without inter-trial intervals. One session consisted of 100 trials (100 reward deliveries) and rats were trained for two sessions per day. The rats were divided into two groups: one group was trained with the FR schedule (n = 3) and the other with the VR schedule (n = 4).

On average, spikes from two or three neurons were recorded simultaneously in each rat. Target neurons were selected based on the criteria described above among the neurons recorded simultaneously and were used for operant conditioning. Only neuron spikes that were recorded > 4 consecutive days were included in data analysis. The time intervals between reward deliveries—not including the 1 s ingestion period—were used as the values of "performance" of operant conditioning. Only the performance values obtained at 100% of threshold were used for comparison between the FR and VR schedules.

Among all data, a review of the video recordings of the rats' bodily movements provided no clear evidence of specific behavioral changes during neuronal operant conditioning. Data of fine bodily movements detected by the three-axis accelerometer from three of the rats also provided no clear behavioral changes prior to reward delivery during the neuronal operant conditioning (Fig. 3). Comparison of the effects of the reinforcement schedules revealed that the performance of each neuron was enhanced in the FR (ANOVA, $F(3,8) = 5.71$, $p < 0.05$) but not in the VR schedule (ANOVA, $F(312) = .21$, n.s.) (Fig. 4A). Compared with the first day, a statistically significant enhancement in performance was observed on the final day of operant conditioning in the FR schedule (t-test, $t(4) = 3.08$, $p < 0.05$). However, there was no significant difference between the performances on the first and the last days in the VR schedule (t-test, $t(6) = 0.17$, n.s.). These results suggest that during the operant conditionings, the rats in the FR schedule learned to

Please cite this article in press as: Song, K., et al., Reinforcement schedules differentially affect learning in neuronal operant conditioning in rats. Neurosci. Res. (2019), https://doi.org/10.1016/j.neures.2019.04.003

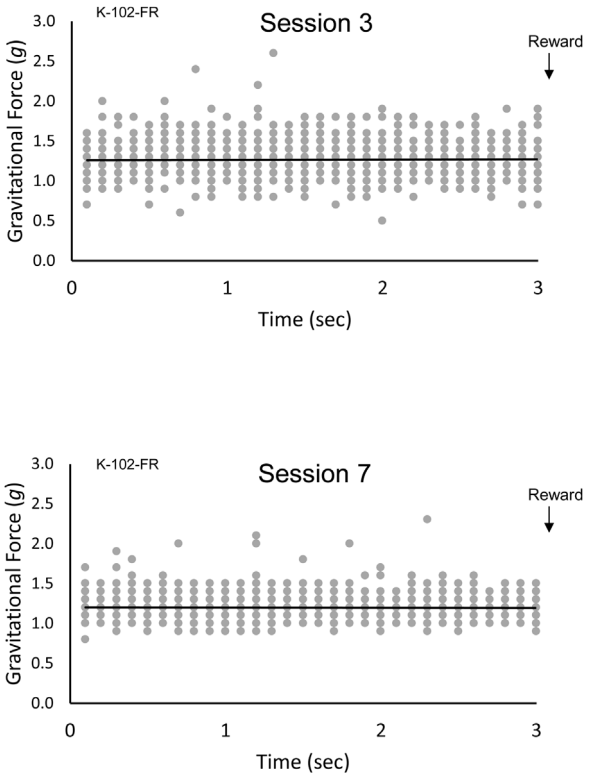

Fig. 3. An example of bodily movements detected by the three-axis accelerometer from one rat during the second day (Session 3, top) and the last day (Session 7, bottom) of the neuronal operant conditioning using FR schedule.
The vertical axis of each plot indicates signal vector magnitude of 3 axes (x, y, z) in gravitational force (Robert et al., 2009). The horizontal axis indicates time in second just prior to reward delivery. The dots represent data of all trials in each session (100 trials) and thus one dot sometimes includes data of several trials. The horizontal straight lines represent results of linear regression analysis to check specific changes in the level of movements in the rat. Session 3 in the second day was used for the data in the early period of conditioning, because the rats were not fully accustomed to the experimental procedure in the first day and had a possibility to show irregular and unusual movements.

enhance their neuron firing rates to obtain rewards; however, the rats reinforced by the VR schedule exhibited no such learning until the last day.

A two-way repeated measures ANOVA demonstrated only marginal significance in the main effect of reinforcement schedules ($F(1,5) = 4.98$, $p = 0.076$) as well as in interaction between reinforcement schedules and training days ($F(1.94,9.72) = 3.17$, $p = 0.088$) in the total period of training. However, the Bonferroni post-hoc analysis indicated that the performance in the FR schedule was significantly higher compared with the VR schedule ($p < 0.05$) on the last day of conditioning (Fig. 4B).

Neuronal operant conditioning with successive FR-VR schedules was also performed to test the effect of different reinforcement schedules on neuronal activity from the same neurons. The FR schedule was applied during the first two days and the VR schedule during the last two days ($n = 2$). Simple linear regression analysis of the FR, VR, and FR-VR data (Fig. 4C) demonstrated that the recorded neurons in the FR-VR schedules improved their performance ($y = -2.6835x + 25.168$, $R^2 = 0.645$, $p < 0.05$) (FR-VR in Fig. 2C), which was similar in FR schedule ($y = -2.3468x + 23.243$, $R^2 = 0.68$, $p < 0.001$). This result was noticeable compared with the poor performance in the VR schedule alone ($y = 0.1022x + 20.126$, $R^2 = 0.002$, n.s.) (VR in Fig. 4C), suggesting that the VR schedule may lead to successful operant conditioning of neuronal activity by being combined with different reinforcement schedules such as the FR.

Most previous studies investigating neuronal operant conditioning used a continuous reinforcement schedule (CRF) or simple intermittent reinforcement schedules in discrete type situations (Arduin et al., 2013; Engelhard et al., 2013; Sakurai and Takahashi, 2013). This study was the first to investigate the effect of basic and different reinforcement schedules in the free operant type situation on neuronal operant conditioning using long-term recording of single neuronal spiking for four days of training. The result demonstrated that neuron activity was successfully conditioned in the FR schedule but not in the VR schedule when they were applied individually (Fig. 4A). This was unexpected because the animals' responses can be easily conditioned in both schedules, and the conditioned response rates are generally higher in VR than in FR schedules in behavioral operant conditioning (Reynolds, 1975). However, the rats could be conditioned in the FR-VR schedules (Fig. 4C), meaning that the training in the FR schedule could "shape" rat's responses applicable in the VR schedule. Therefore, the results also conclude that neuronal activity was successfully conditioned in the FR schedule without shaping procedures, but in the VR schedule it could be conditioned only with the shaping procedure.

On the basis of observation by the video camera and detection of fine movements by the three-axis accelerometer (Fig. 3), we concluded that there was no specific behavioral change during the neuronal operant conditioning. This indicates that motor cortical neurons could be conditioned to increase their firing rates with no increment of bodily movements. The key mechanism to explain such discrepancy of motor cortical activity and bodily movements was the learning-dependent plastic connections between motor cortical neurons and body muscles, formerly suggested in Fetz and Cheney (1987) and Chapin et al. (1999).

While neuronal and behavioral operant conditioning appear to share basic features in underlying assumptions and outcomes, some key features, such as reinforcement schedules, have not been studied in detail in neuronal operant conditioning. Although the mechanisms of such key features are currently unclear, it is cer-

Please cite this article in press as: Song, K., et al., Reinforcement schedules differentially affect learning in neuronal operant conditioning in rats. Neurosci. Res. (2019), https://doi.org/10.1016/j.neures.2019.04.003

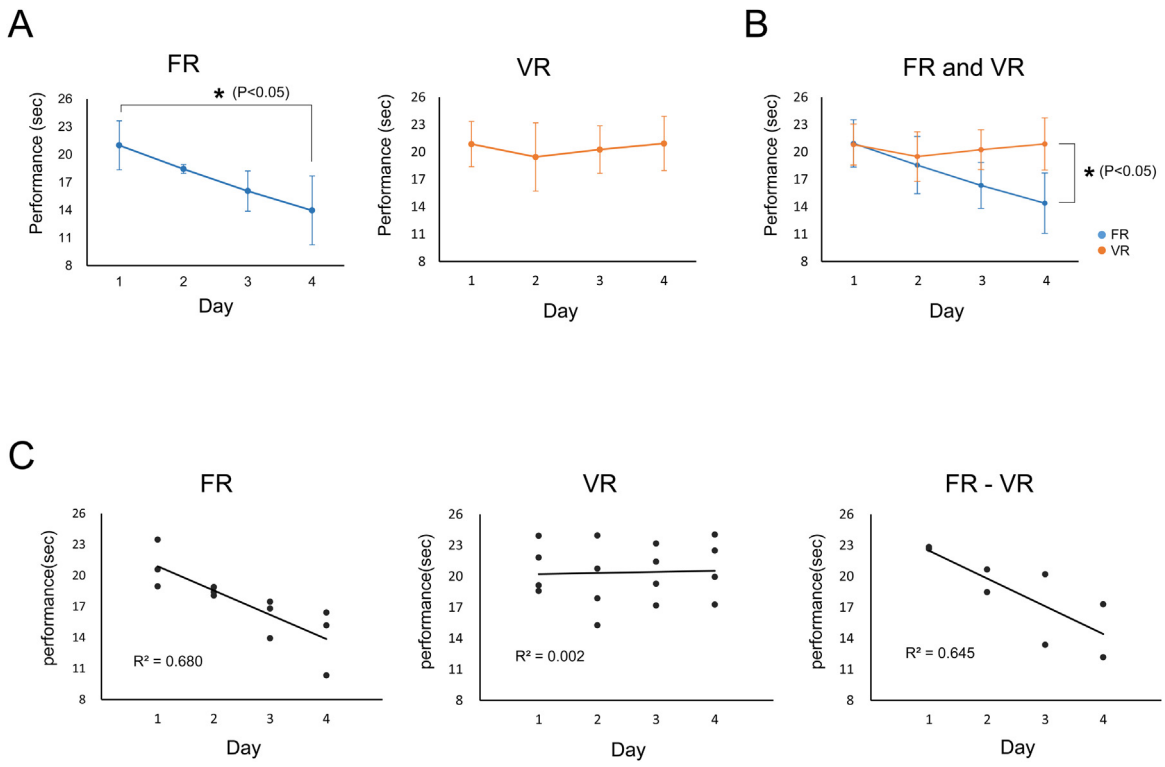

Fig. 4. The effect of different reinforcement schedules on the rats' performance across four successive training days.
The rats were trained for two sessions per day, and the data from two sessions in one day were used for statistical analysis. The vertical axes indicate performance, interval times taken for the rats to obtain a reward in each trial and, therefore, the lower value means the better performance. (A) The performance of rats in the fixed ratio (FR, left) and variable ratio (VR, right) schedules during neuronal operant conditioning. (B) The interaction between reinforcement schedules and training days. (C) Linear regression analysis for FR alone (left), VR alone (middle), and successive FR-VR (right) schedules.

tain that the principles of operant conditioning of neuronal activity are different from those of behavioral responses. Moving to the practical perspective for the development of BMIs, the FR schedule of reinforcement may be more appropriate for training to control external devices using BMIs.

Although this short paper describes only the learning processes in neuronal operant conditioning with different schedules of reinforcement, changes in firing rate and synchrony of neurons during conditioning should be examined in detail to observe dynamic changes in neuronal functions in operant conditioning (e.g., Sakurai and Takahashi, 2013). We plan to report the results of such an examination in the near future.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Numbers 16H02061 and 18H05088 to YS. The authors thank Dr. Shoko Yuki for her support of data analysis.

References

Andersen, R., Hwang, E., Mulliken, G., 2010. Cognitive neural prosthetics. Annu. Rev. Organ. Psychol. Organ. Behav., 1–28.
Arduin, P., Fregnac, Y., Shulz, D.E., Ego-Stengel, V., 2013. "Master" neurons induced by operant conditioning in rat motor cortex during a brain-machine interface task. J. Neurosci. 33, 8308–8320.
Baranauskas, G., 2014. What limits the performance of current invasive brain machine interfaces? Front. Syst. Neurosci. 8, 68.
Chapin, J.K., Moxon, K.A., Markowitz, R.S., Nicolelis, M.A.L., 1999. Real-time control of a robot arm using simultaneously recorded neurons in the motor cortex. Nat. Neurosci. 2, 664–670.
Collinger, J.L., Wodlinger, B., Downey, J.E., Wang, W., Tyler-Kabara, E.C., Weber, D.J., McMorland, A.J.C., Velliste, M., Boninger, M.L., Schwartz, A.B., 2013. High-performance neuroprosthetic control by an individual with tetraplegia. Lancet 381, 557–564.
Donoghue, J.P., Nurmikko, A., Black, M., Hochberg, L.R., 2007. Assistive technology and robotic control using motor cortex ensemble-based neural interface systems in humans with tetraplegia. J. Physiol. (Paris) 579, 603–611.
Engelhard, B., Ozeri, N., Israel, Z., Bergman, H., Vaadia, E., 2013. Inducing gamma oscillations and precise spike synchrony by operant conditioning via brain-machine interface. Neuron 77, 361–375.
Ethier, C., Oby, E.R., Bauman, M.J., Miller, L.E., 2012. Restoration of grasp following paralysis through brain-controlled stimulation of muscles. Nature 485, 368–371.
Fetz, E.E., 1969. Operant conditioning of cortical unit activity. Science 163, 955–958.
Fetz, E.E., 2007. Volitional control of neural activity : implications for brain – computer interfaces. Society 579, 571–579.
Fetz, E.E., Cheney, P.D., 1987. Functional relations of primate motor cortex cells and muscles: fixed and flexible. In: Motor Areas of Cerebral Cortex, Ciba Foundation Symposium 132. Willey, Chichester, pp. 98–124.
Hwang, E., Bailey, P.M., Andersen, R., 2013. Volitional control of neural activity relies on the natural motor repertoire. Curr. Biol. 23, 353–361.
Koralek, A.C., Jin, X., Long, J.D., Costa, R.M., Carmena, J.M., 2012. Corticostriatal plasticity is necessary for learning intentional neuroprosthetic skills. Nature 483, 331–335.
Lebedev, M.A., 2014. Brain-machine interfaces: an overview. Transl. Neurosci. 5, 99–110.
Lebedev, M.A., Nicolelis, M.A.L., 2011. Toward a whole-body neuroprosthetic. Prog. Brain Res. 194, 47–60.
Mandonnet, E., Duffau, H., 2014. Understanding entangled cerebral networks: a prerequisite for restoring brain function with brain-computer interfaces. Front. Syst. Neurosci. 8, 6–11.
Nicolelis, M.A.L., 2011. Mind out of body. Sci. Am. 304, 80–83.
Nicolelis, M.A.L., Lebedev, M.A., 2009. Principles of neural ensemble physiology underlying the operation of brain-machine interfaces. Nat. Rev. Neurosci. 10, 530–540.
Reynolds, G.S., 1975. A Primer of Operant Conditioning. Scott Foresman, Glenview.
Robert, B., White, B.J., Renter, D.G., Larson, R.L., 2009. Evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle. Comput. Electron. Agric. 67, 80–84.
Sakurai, Y., 2014. Brain-machine interfaces can accelerate clarification of the principal mysteries and real plasticity of the brain. Front. Syst. Neurosci. 8, 104.
Sakurai, Y., Song, K., 2016. Neural operant conditioning as a core mechanism of brain-machine interface control. Technologies 4, 26.
Sakurai, Y., Takahashi, S., 2013. Conditioned enhancement of firing rates and synchrony of hippocampal neurons and firing rates of motor cortical neurons in rats. Eur. J. Neurosci. 37, 623–639.

Please cite this article in press as: Song, K., et al., Reinforcement schedules differentially affect learning in neuronal operant conditioning in rats. Neurosci. Res. (2019), https://doi.org/10.1016/j.neures.2019.04.003

Sakurai, Y., Song, K., Tachibana, S., Takahashi, S., 2014. Volitional enhancement of firing synchrony and oscillation by neuronal operant conditioning: interaction with neurorehabilitation and brain-machine interface. Front. Syst. Neurosci. 8, 11.
Takahashi, S., Sakurai, Y., 2005. Real-time and automatic sorting of multi-neuronal activity for sub-millisecond interactions in vivo. Neuroscience 134, 301–315.
Takahashi, S., Sakurai, Y., 2007. Coding of spatial information by soma and dendrite of pyramidal cells in the hippocampal CA1 of behaving rats. Eur. J. Neurosci. 26, 2033–2045.
Takahashi, S., Sakurai, Y., 2009a. Sub-millisecond firing synchrony of closely neighboring pyramidal neurons in hippocampal CA1 of rats during delayed non-matching to sample task. Front. Neural Circuits 3, 9.
Takahashi, S., Sakurai, Y., 2009b. Information in small neuronal ensemble activity in the hippocampal CA1 during delayed non-matching to sample performance in rats. BMC Neurosci. 10, 115.
Velliste, M., McMorland, A.J.C., Kennedy, S.D., Whitford, A.S., Sohn, J.W., Schwartz, A.B., 2014. Motor cortical correlates of arm resting in the context of a reaching task and implications for prosthetic control. J. Neurosci. 34, 6011–6022.

Please cite this article in press as: Song, K., et al., Reinforcement schedules differentially affect learning in neuronal operant conditioning in rats. Neurosci. Res. (2019), https://doi.org/10.1016/j.neures.2019.04.003