Simple Associative Learning Accounts for the Complex Dynamics of Operant Extinction

José R. Donoso (jose.donoso@rub.de)
Institute for Neural Computation, Ruhr-Universität Bochum, Universitätstr. 150
Bochum, 44801 Germany

Zhiyin Lederer (zhiyin.li@rub.de)
Institute for Neural Computation, Ruhr-Universität Bochum, Universitätstr. 150
Bochum, 44801 Germany

Julian Packheiser (julian.packheiser@rub.de)
Department of Biopsychology, Ruhr-Univeristät Bochum, Universitätstr. 150
Bochum, 44801 Germany

Roland Pusch (roland.pusch@rub.de)
Department of Biopsychology, Ruhr-Univeristät Bochum, Universitätstr. 150
Bochum, 44801 Germany

Thomas Walther (thomas.walther@rub.de)
Institute for Neural Computation, Ruhr-Universität Bochum, Universitätstr. 150
Bochum, 44801 Germany

Onur Güntürkün (onur.guentuerkuen@rub.de)
Department of Biopsychology, Ruhr-Univeristät Bochum, Universitätstr. 150
Bochum, 44801 Germany

Sen Cheng (sen.cheng@rub.de)
Institute for Neural Computation, Ruhr-Universität Bochum, Universitätstr. 150
Bochum, 44801 Germany

Abstract:

Extinction Learning (EL), the ability of changing a previously acquired behaviour as a result of altered reinforcement contingencies, is essential for a successful adaptation to a changing environment. EL is specific to the context in which it occurred, as evidenced by the so-called renewal of the extinguished behaviour after returning to the original context where conditioning took place (ABA renewal). Here, we analyse single learning curves obtained from pigeons performing an operant conditioning task within an ABA paradigm. The curves exhibit a stunning diversity both across subjects and within a subject across sessions. We find that a computational model with simple sensorimotor associations and a winner-takes-all decision process can surprisingly account for most of the peculiar features in the behaviour. Our model suggests that the complexity of behaviour stems from the history of context and rewarded responses within and across sessions. In conclusion, our work demonstrates how studying the dynamics of learning can reveal previously unappreciated nuances in the behaviour and how even simple models can generate complex, apparently purposeful behaviour.

Keywords: extinction learning; operant conditioning; associative models.

Experimental Paradigm

Pigeons underwent several learning sessions, each of which consisted of three phases. In the initial acquisition phase, under context A, animals had to learn to associate two session-unique novel visual stimuli with either a left, or a right-peck response, where food was delivered after every correct choice (See Fig. 1A). In the subsequent extinction phase, under either a context B1 or B2, one of the novel stimulus-response associations was no longer rewarded. Once pigeons stopped choosing the associated behavior, the test phase began with a return to the context A, but the extinguished association continued to go unrewarded. Throughout the session, pigeons were additionally presented with two familiar control stimuli, for which the reward contingency remained unchanged (for details, see Packheiser, Pusch & Güntürkün, 2019).

Dynamics of Learning Exhibits a Broad Diversity

In the literature of extinction learning, changes in behaviour are typically quantified by comparing post to pre-training blocks (e.g., Todd, 2013; Trask, Thrailkill & Bouton, 2017). To this end, data are pooled and averaged across many sessions and subjects. Such analysis obscures not only differences across subjects but also interesting aspects of the dynamics of learning within and across sessions. To uncover such hidden dynamics, we analysed single learning curves (Gallistel, Fairhurst & Balsam, 2004) and assessed their variability across sessions and across subjects. Learning curves exhibited a wide repertoire of dynamics, expressing combinations of different features such as persistent selection of unrewarded responses, abrupt transitions of choice upon the onset of the extinction context, and absence (and later reappearance) of the renewal effect, among others. These complex behaviors appear to be strategic at first glance and to rely on higher-order cognitive functions.

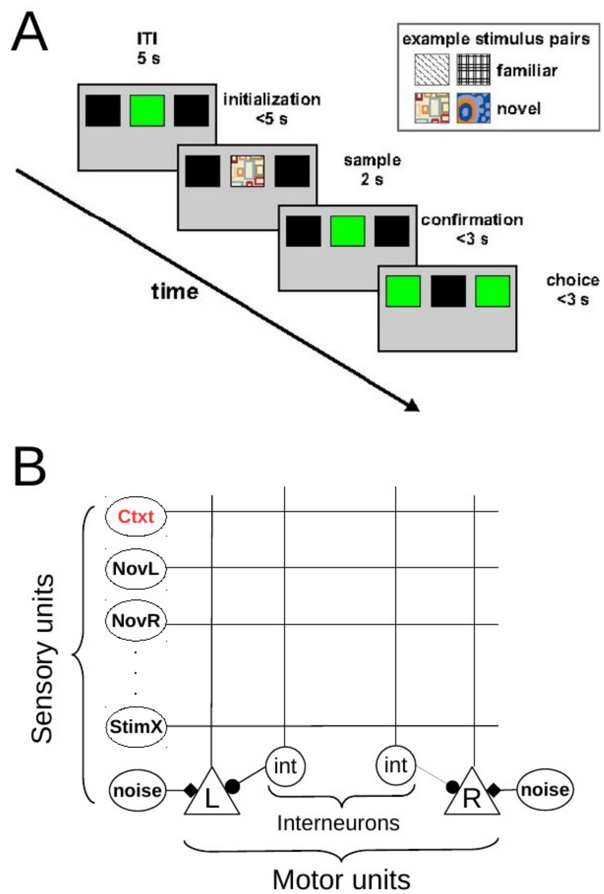

Figure 1: (A) Time-course of a single trial (adapted from Starosta, Stüttgen & Güntürkün, 2014). (B) Associative network model.

Associative Learning can Generate Complex Learning Dynamics

In order to asses to which extent associative models could account for the seemingly top-down-controlled behavior of the pigeons, we implemented an associative network model (Fig.1B). In this model, sensory units (ovals) signaling the presence of specific stimuli can establish direct excitatory connections with the motor units (triangles) mediating the left (L) and right (R) responses. Sensory units can also establish connections with interneurons (circles) that can inhibit a specific motor response. Connections to motor units and interneurons are reinforced in a Hebbian fashion any time a reward is delivered or withdrawn, respectively. The model selects the response corresponding to the motor unit with the highest activation. When both motor units are inhibited below their threshold of activation, no response is selected (omission).

A key aspect of our model is that the context is treated as just another stimulus and therefore, it can directly establish excitatory and inhibitory associations with specific responses (Rescorla, 1993, 1997). This property enables the context to generate imbalances in the net inputs to motor units, which can explain the emergence of persistent unrewarded behaviour during the extinction phase. Furthermore, the same property enables the network to carry previously established associations with a specific context to the next session, leading to a wide diversity of learning curves across sessions.

Conclusion

Here, we report the stunning diversity of learning curves exhibited by pigeons in an operant conditioning task, and offer a simple model that can account for such diversity. Our data and model gives support to the hypothesis that context can establish direct excitatory and inhibitory associations with specific responses (Rescorla 1993, 1997), as opposed to the gating mechanism proposed by the 'occasion setting' hypothesis (Bouton, 2004).

Acknowledgments

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number project number 316803389 – SFB 1280, projects F01, A01 and A14.

References

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory 11, 485-494.*

Gallistel, C. R., Fairhurst, S. & Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences, 101, 13124-13131*.

Packheiser, J., Güntürkün, O., & Pusch, R. (2019). Renewal of extinguished behavior in pigeons (Columbia livia) does not require memory consolidation of acquisition or extinction in one-day free-operant appetitive conditioning paradigms. *Behavioral Brain Research, 111947*.

Rescorla, R. A. (1993). Inhibitory associations between S and R in extinction. *Animal Learning & Behavior 21, 327-336.*

Rescorla, R. A. (1997). Response inhibition in extinction. *The quarterly Journal of Experimental Psychology 50B, 238-252.*

Starosta, S., Stüttgen, M. C. & Güntürkün, O. (2014). Recording single neuron's action potentials from freely moving pigeons across three stages of learning. *Journal of Visualized Experiments 88, e51283.*

Todd, T. P. (2013). Mechanisms of renewal after the extinction of instrumental behavior. J*ournal of Experimental Psychology: Animal Behavior Processes 39, 193-207.*

Trask, S., Thrailkill, E. A. & Bouton, M. E. (2017). Occasion setting, inhibition, and the contextual control of extinction in Pavlovian and instrumental (operant) learning. *Behavioral Processes 137, 64-72.*