Choice behavior in dynamic Random-Interval Random-Ratio schedules of reinforcement

Manuel Villarreal; Carlos Velazquez; Arturo Bouzas

Universidad Nacional Autónoma de México (UNAM); Facultad de Psicología

Lab 25; Supported by Proyecto PAPIIT IG100818

Introduction

From the perspective of animal learning, the study of adaptation to volatile environments has been carried out using programs where the probability of being reinforced depends on the time elapsed since the last time a reinforcer was obtained from the alternative (Random Interval, RI). When an organism has to choose between two options with this structure but different values (concurrent schedules), the allocation of behavior to each alternative has little impact on the total amount of reinforcers that can be obtained. Additionally, the ratio of responses tends to match the ratio of obtained reinforcers, a result known as Matching. A second class of reinforcement functions are known as Random Ratio (RR), in this case, each response at the alternative has a constant probability of being reinforced, which is similar to a bandit problem. If there are two alternatives one of which depends on time elapsed and the other on behavior, then the amount of reinforcement that an organism can obtain depends on how behavior to each alternative is allocated. For this experiment, the objective was to study how organisms adapted their behavior when the values of two different reinforcement rules (RR-RI) change frequently and abruptly within a session.

Method

Pigeons could choose continuously between two response options. In the first one, a response was always associated with a probability of obtaining a reinforcer (RR). In the second one, there was a constant probability of a reinforcer being available each second (RI). Within a single session, every ten reinforcers, one of this probabilities changed while the other was held constant. Ten different pairs of probabilities were presented. For five of these pairs, the RR schedule was held fixed at 30 responses while varying the values of the RI schedule; the possible values for this program were 7.5, 15, 30, 60 and 120s; after ten reinforcers were received from either program, a new value was drawn at random and without replacement. For the remaining five pairs, the value of the RI schedule was held fixed at 60s, while the possible values of the RR schedule were 15, 30, 45, 60 and 120 responses; which were also sampled randomly and without replacement each experimental session. The analysis presented here use the data of the last 60 sessions out of 110.

Results

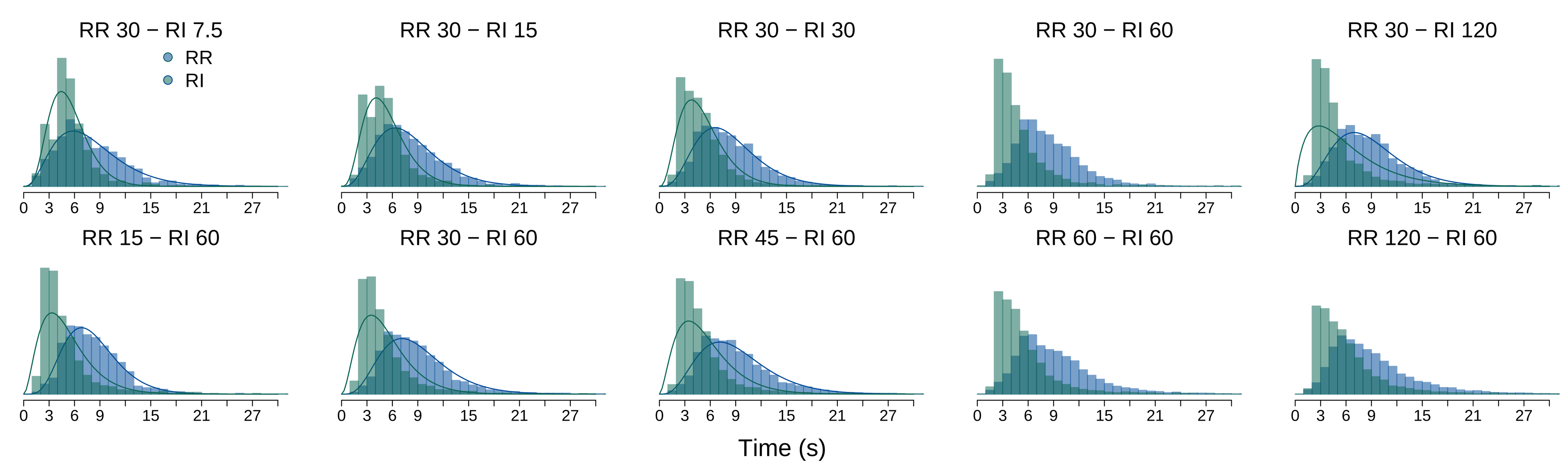

Figure: Visit length in seconds at each response option when the visit was not interrupted by a reinforcer. Curves show the estimated density when the distributions are different.

Results

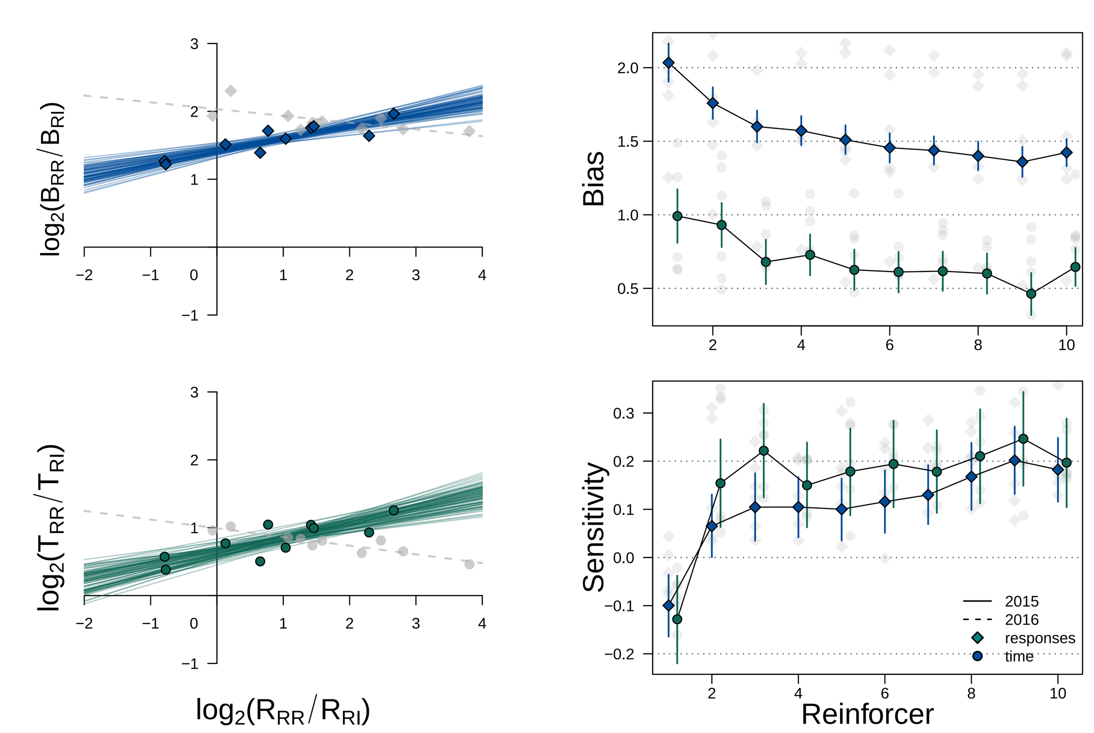

Figure: Generalized Matching: logarithm of responses (top) and time (bottom) as a function of the logarithm of reinforcers before the first (gray) and tenth reinforcer. In the second column we present the mean and 95% High Posterior Density Interval for the intercept (top, Bias) and slope (bottom, Sensitivity), of a linear regression before each reinforcer with the estimates for responses in blue and time in green, gray marks represent mean of individual pigeons.

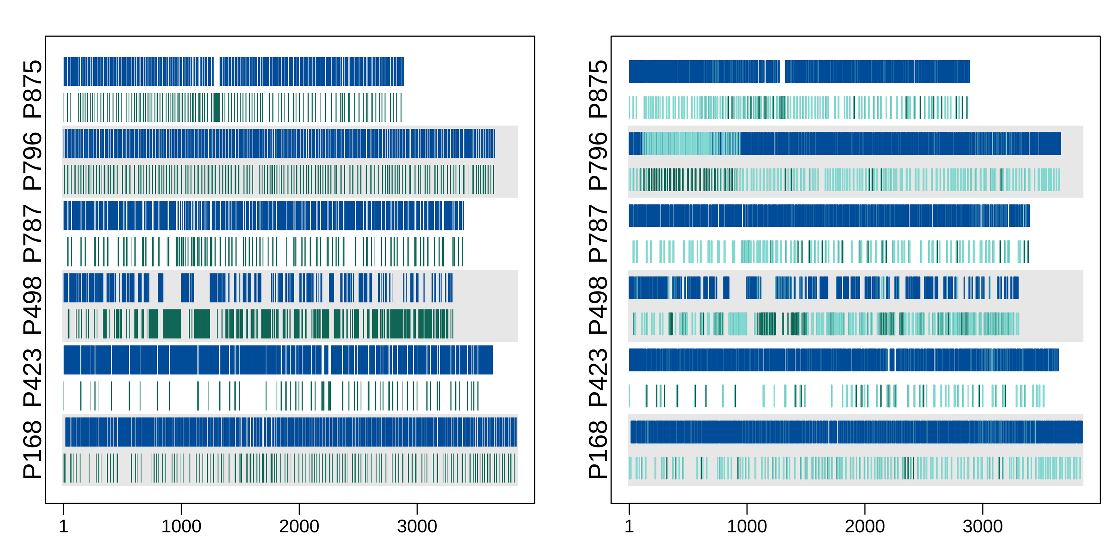

Figure: Response by response data in the last session. green bars represent responses to RI, blue lines represent responses to RR. Light blue lines in the second pannel represent incorrect prediction by the value model.

Bayesian Modeling

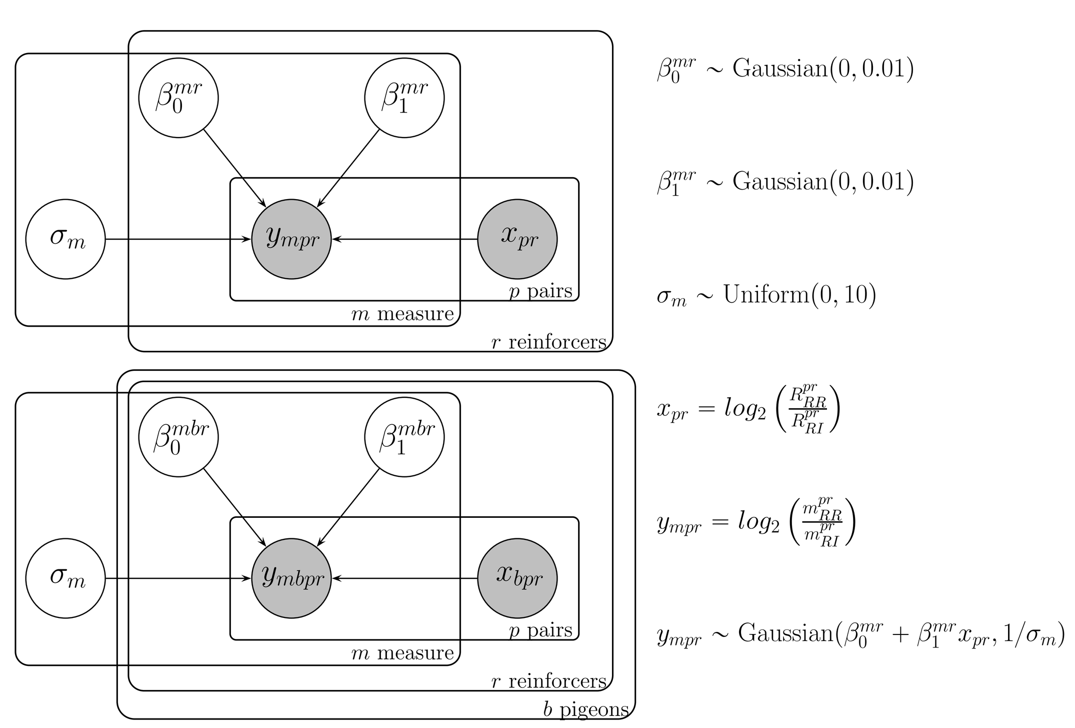

Figure: Generalized Matching: Bayesian graphical model for the collapsed data (top) and individual data (bottom).

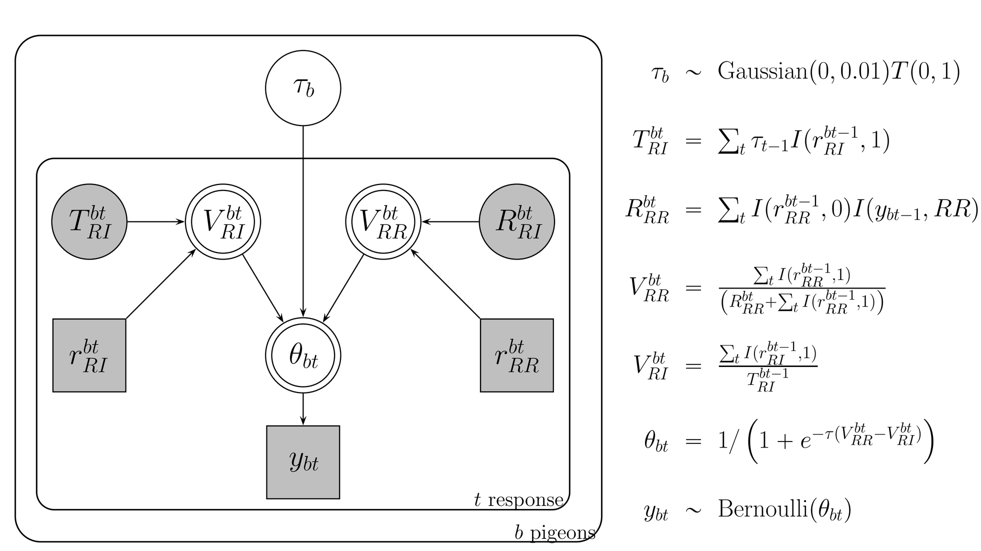

Figure: Dynamic model used to predict behavior in the last session were value for both programs is defined as the probability of being reinforced. The probability of choosing the RR schedule is generated using a logistic function.

Discussion

1) Visits
   Histograms in the first figure shows that the amount of time in seconds that a pigeon spends in each option depends on the reinforcement rule, with larger times spend on the Random Ratio alternative. This suggest that pigeons are able to discriminate between different reinforcement schedules and is consistent with an optimization rule.
2) Matching
   The panels in the right of the second figure show how pigeons allocated their behavior as a function of the log ratio of reinforcers. There was a substantial bias in favor of the RR option, nevertheless there was more responding to the RI option that was needed to collect all the available reinforcers, generating a substantial undermatching. Additionally, the panels on the right show that the bias towards the RR decreased as more reinforcers were obtained, indicating an increasing control by the expected time to reinforcement on the RI option. The impact of reinfocement ratios increased after the first reinforcer and remained relatively constant afterwards, indication that the adaptation to a new set of probabilities required only one or two reinforcers.
3) Dynamic model
   The last figure shows the behavior of all the pigeons in the last experimental session. As can be seen from the second panel of the figure, responses to the RI schedule are hard to predict for a model that uses probability of reinforcement as a value function. This can be expected given that the pigeons respond more often to the RI alternative from what can be expected from an optimization perspective.

References

Herrnstein, R. J., & Heyman, G. M. (1979). Is Matching Compatible with Reinforcement Maximization on Concurrent Variable Interval, Variable Ratio? Journal of the Experimental Analysis of Behavior, 31(2), 209-223.

Villarreal, M., Velázquez, C., Baroja, J. L., Segura, A., Bouzas, A., & Lee, M. D. (2019). Bayesian Methods Applied to the Generalized Matching Law. Journal of the Experimental Analysis of Behavior, 111(2), 252-273.

Bell, M. C., & Baum, W. M. (2017). Concurrent Variable-Interval Variable-Ratio Schedules in a Dynamic Choice Environment. Journal of the Experimental Analysis of Behavior, 108(3), 367-397.

Contact Information

Laboratorio 25: bouzaslab25.com

Manuel Villarreal: jesus.mvu@gmail.com

Carlos Velazquez: carlos.unamlab25@gmail.com