

Matching without learning

Pier-Olivier Caron

Adaptive Behavior

1–7

© The Author(s) 2015

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/1059712315590484

adb.sagepub.com



Abstract

Matching behavior is a phenomenon describing response rate ratios of an organism as a function of their associated reinforcer rate ratios. The generalized matching law (GML), its quantitative formulation, has been frequently found to explain over 80% of the variance in concurrent reinforcement schedules. However, a previous paper found by means of Monte Carlo simulations that matching behavior could be due to environmental constraints on behavior rather than a mere decision-making process. The purpose of the current study is to systemically investigate the influence of constraints induced by concurrent schedules of reinforcement. A Monte Carlo simulation was carried out. Results showed that the GML reached much better explained variances with real (and artificial) organisms than the current simulated results. Thus, a learning process seems partly necessary to generate matching behavior. According to the current findings, concurrent reinforcement schedules clearly induced a quantitative dependency between behavior rates and reinforcer rates. The simulation demonstrates that matching behavior is not only a consequence of a behavioral (decision-making) process, but of environmental conditions also.

Keywords

Matching law, Monte-Carlo simulations, operant conditioning, matching behavior, learning

1 Introduction

In a concurrent reinforcement schedule (much like a two-armed bandit), much attention has been given to the molar relation concerning an organism's responses as a function of reinforcers, e.g., the matching law (Baum, 1974). The most common version of the law in experimental behavior analysis is the generalized matching law (GML) (Baum, 1974), which states that from a session-by-session standpoint (or a molar perspective) the responses ratio of an organism should conform to the reinforcer ratios. This relationship is represented by the following equation:

$$\log(B_1/B_2) = a \log(R_1/R_2) + \log c \quad (1)$$

where B_1 and B_2 correspond to response rates, R_1 and R_2 correspond to reinforcer rates, parameter a designates sensitivity, and the parameter c designates the bias. After a short learning period with a given reinforcer ratio, the organism reaches the steady state predicted by Equation 1, and if the reinforcers ratio changes, the organism's behaviors will adjust accordingly. This phenomenon is sometimes called matching behavior. Even though the matching equation has been extensively studied and found to hold in several experimental and natural studies (Davison & McCarthy, 1988; McDowell, 2013a; Wearden & Burgess, 1982),

the underlying processes by which organisms reach the steady state has yet to be understood.

Several computational models have been developed over the years to account for the molecular performance (a response-by-response perspective) of organisms in concurrent reinforcement schedules, i.e., matching behavior. These include a genetic algorithm of selection by consequences (McDowell, 2013b), a Bayesian decision-making algorithm (Saito, Katahira, Okanoya & Okada, 2014), and many neural network algorithms (Dawson, Dupuis, Spetch & Kelly, 2009; Iigaya & Fusi, 2013; Loewenstein, 2008). Though the computation involved in each of these models is quite different, each is based on the same qualitative mechanism: a quantitative learning process contingent on the distribution of reinforcers (or reinforcement learning; Sutton & Barto, 1998). For convenience, the action selection in concurrent schedules is generally stochastic rather than deterministic (McDowell, 2004; Shimp,

Laboratoire des Sciences Appliquées du Comportement, Département de Psychologie Université du Québec à Montréal, Montreal, Quebec, Canada

Corresponding author:

Pier-Olivier Caron, Laboratoire des Sciences Appliquées du Comportement, Département de Psychologie, Université du Québec à Montréal, C.P. 8888, succ. Centre-ville, Montréal, Quebec, H3C 3P8, Canada.

Email: pocaron19@gmail.com

1992). At every time step t , the algorithm updates the probability of B_i ($i=1, 2$) according to whether a reinforcer has been delivered (which increases the probability) or not (decreases the probability). Thus, all models of matching behavior are related to melioration theory, the first theoretical model to account for the matching law. Melioration theory (Herrnstein, 1997) states that the agent compares the rates of return from possible alternatives and shifts toward the one yielding the highest return. From a computational standpoint, it is the same as stipulating that an action selection is based on reinforcer rate ratios (i.e., the matching law), but according to a molecular perspective. It is not surprising that, despite the algorithmic differences between models, they all reach the same computational goal: to mimic matching behavior and respect the GML's predictions. The phenomenon of interest is built within the algorithm. Sakai and Fukai (2008) even showed that partial maximization, if it includes past choices and other available information of reward, will lead to matching behavior regardless of the mathematical algorithm. Thus, theoretical models with similar computational outcomes cannot be distinguished empirically.

Past models have aimed to explain matching behavior as a behavioral process within the organism through action selection and learning (e.g., McDowell, 2013b), or estimation of expected incomes (e.g., Gallistel, 2005; Saito, et al., 2014). However, none has tried to account for matching behavior through the use of environmental processes. An unconventional way to look at matching behavior is by the means of environmental constraints¹ (properties restraining choices) rather than mere behavioral processes (e.g., decision-making algorithms or reinforcement learning). These constraints are rooted in a feedback system made up of interactions between an organism's behavior and its environment (Baum, 1973). Feedback functions describe how the consequences of behavior feeds back its effect to the organism, and inversely how an organism, through its behavior, produces consequences. In other terms, it conceptualizes the quantitative behavior–environment interaction.

Inherent to reinforcement schedules, feedback systems impose constraints on behaviors and reinforcers. For instance, an organism's responses and their associated reinforcers are systematically constrained by their quantitative interdependency: the number of reinforcers cannot be higher than the number of responses (McDowell & Ansari, 2005). Equation 2 illustrates this mathematical relationship;

$$R_{ij} \leq B_{ij} \quad (2)$$

where variables indicate the number of obtained reinforcers (R) or behaviors (B) from option i at the session j ($j=1, 2, 3, \dots, n$). Equation 2 comes from the basic definition of a contingency of reinforcement (Skinner,

1938); a response must be emitted in order to be reinforced. In fact, few instances of behavior are reinforced through concurrent reinforcement schedules (McDowell & Ansari, 2005). If no behavior is emitted before the reinforcer occurs, then either (a) another response has been reinforced or (b) it is not a reinforcer.

A probability density function offers a simple way to see the influence of this constraint. The upper panels of Figure 1 show the empirical ($j=100,000$) probability density function of behaviors (left) and reinforcers (right) sampled from a uniform distribution, but where R_{ij} is constrained by the upper bound B_{ij} . The influence of Equation 2 can be seen in the upper right panel: the distribution is exponential rather than uniform. The two panels in the middle row show the logarithm of the ratio of these distributions, which appears to be Gaussian. Finally, the bottom panel presents both distributions as analyzed with the GML (the x -axis shows the reinforcer rate ratio and y -axis shows the response rate ratio). A linear trend is visible to the naked eye as well as statistically ($r^2 = .50$, $a = .50$ and $\log c = .00$). Matching behavior would have occurred in a biological agent.

Equation 2 is not the only constraint imposed by a concurrent reinforcement schedule. Another relevant constraint is the theoretical maximum number of possible responses (B_t) or reinforcers (R_t) such as:

$$B_1 + B_2 = B_t \quad (3)$$

or

$$R_1 + R_2 = R_t \quad (4)$$

which are computationally similar. Equations 3 and 4 have been discussed in matching theory (Herrnstein, 1997; Rachlin, Battalio, Kagel & Green, 1981). Their theoretical implications for living agents go beyond the scope of this paper. It suffices to say that all experimental and natural studies (within the experimental analysis of behavior) with organisms end when a given time has elapsed (e.g., 1 h) or a given amount of reinforcers have been distributed (e.g., 60), whichever comes first. It is impossible to observe a living organism (and artificial ones either) emitting an infinite number of responses. The same analysis as depicted in Figure 1 was carried out using Equation 3 and 4. The visual representations of results are quite similar. However, an ordinary least squares regression revealed a stronger relationship between the reinforcer rate ratio and response rate ratio ($r^2 = .62$, $a = .62$ and $\log c = .00$).

These preliminary results demonstrate that, regardless of the effect of the reinforcer, fitting the GML to behavioral data should produce spurious correlations (i.e., systematically positive non-null correlations regardless of sampling errors). The complexity of matching behavior can be conceptualized largely as a reflection of the complexity of the environment and, more specifically, of how behaviors and reinforcers are

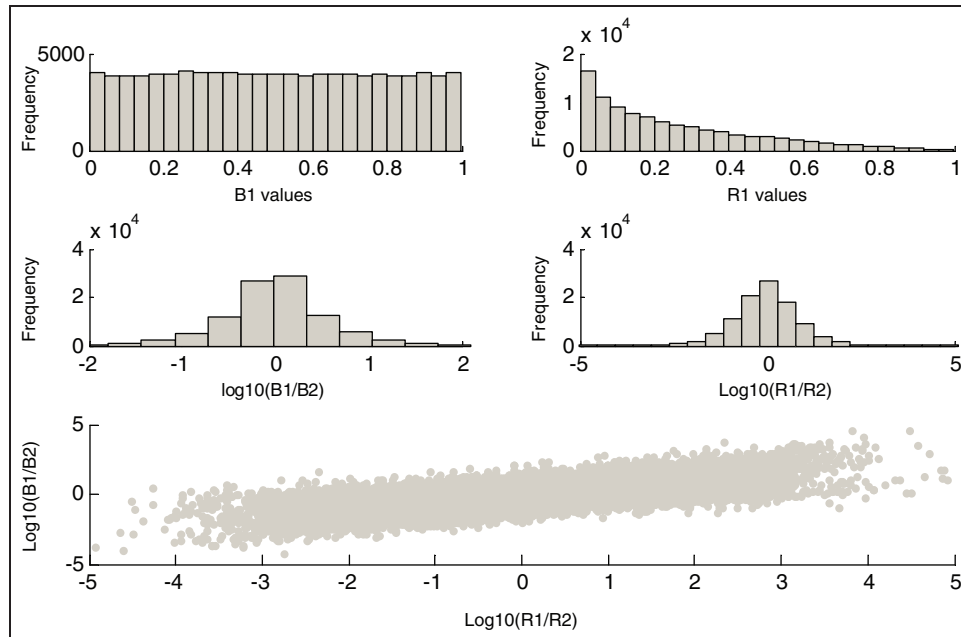


Figure 1. The upper panels show the empirical ($j=100,000$) probability density function of behaviors (left) and reinforcers (right). Behaviors were sampled from a uniform distribution, $[0, 1]$. Data from the upper right panel was sampled according to Equation 2, i.e., where B_{ij} was the upper bound of R_{ij} , $[0, B_{ij}]$. The middle panels show the logarithm of the ratio of two samples (same as upper panels) responses (left) and reinforcer (right). The bottom panel shows data as fitted according to the generalized matching law (GML) (x -axis as the reinforcer rate ratio and y -axis as the response rate ratio). A linear trend is observed ($r^2=.50$, $a=.50$ and $\log c = .00$; see Equation 1 for parameters).

observed. Thus, matching is not necessarily achieved by any formal decision-making algorithm, but could be induced by concurrent reinforcement schedules (Caron, 2014). The equilibrium stated by Equation 1 can be seen as not the mere product of a learning process, but by the mathematical properties that relate behaviors to reinforcers.

Despite the constraints induced by concurrent schedules of reinforcement having a substantial effect, no investigation of their influence has been carried out. To address this issue, the current study simulates the constraints of Equation 2, which stipulates that the amount of reinforcers cannot be higher than the number of responses, and Equation 3, which constrains the theoretical number of possible responses by session. The objective is to assess their influences on the GML. It offers a new perspective of the GML as a product of the environment (as will be modeled in the remaining of the study) rather than a behavioral process. In order to mimic the real-life observation procedure of matching behavior, both equations will be tested with a Monte-Carlo simulation method.

2 Numerical simulation I

In order to investigate the influence of the first constraint on matching behavior, a Monte Carlo simulation was carried out. The simulation will serve to replicate that of Caron (2014), which found that the

GML could account for 47% of the variance when Equation 2 was implemented.² Though simulations by Caron (2014) were limited to 16 conditions, the current simulation will include 50 conditions with a larger range of parameters. Therefore, the purpose of the first simulation is twofold: (1) to replicate Caron's results, and (2) to test additional conditions (sample size and maximum bound of responses, see below for details).

A Monte Carlo approach was used to assess the effect of implementing Equation 2, i.e., the amount of reinforcers is always equal to less than the number of responses. The simulation was carried out in Matlab (2012a) and was as follows:

- (1) Generate two samples of observed behavior (B_i) with length n from a discrete uniform distribution $(1, B_i)$;
- (2) Generate two samples of observed reinforcers from a discrete uniform distribution $(1, B_{ij})$ such as every $R_{ij} \leq B_{ij}$;
- (3) Fit Equation 1 to the data set using ordinary least squares correlations and record a , $\log c$ and the r^2 ;
- (4) Repeat steps 1, 2, and 3 a total of 10,000 times.

Number of sessions ($j=1, 2, 3, \dots, n$), or how many times an organism is observed (i.e., the sample size),³ and B_i (maximum bound of the number of responses) was varied from 10, 20, 30, 40, and 50, and 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100, respectively.

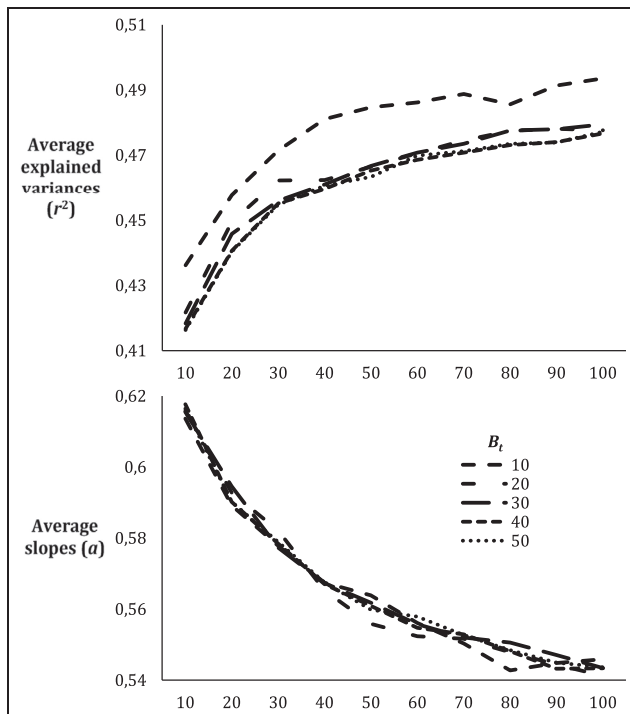


Figure 2. Results from simulation 1. Top panel depicts the average explained variances of the generalized matching law (GML) according to the number of sessions (n) (abscissa) and the upper bound of the uniform distribution (B_t) (line style). The expected explained variance reaches an asymptote approximately at $n=30$, that finally reaches .47 at $n=80$. There were substantial influences of B_t , especially on the ($n=10$) conditions which seems more sensitive to sample variations. Bottom panel presents the average slopes according to the same conditions (abscissa and line style). Average slope decelerated pattern is relatively orderly.

2.1 Results

The current Monte Carlo simulations revealed that the GML explained on average 47% of the variance when Equation 2 was implemented, supporting preliminary work by Caron (2014). Top panel of Figure 2 shows that when $n=10$, the expected explained variance is .41, an asymptote is reached at approximately $n=30$ or 50, the variance finally reaching .47 at $n=80$. There were substantial influences of B_t on the explained variance, especially for the ($n=10$) conditions, which showed higher average explained variances than others B_t values. The bottom panel of Figure 2 depicts the influence of B_t and n on the slope (a) of the GML. Average slope patterns are relatively orderly (at the third decimal). The results show that expected sensitivity slightly decreases as the number of sessions increase. An asymptote is reached at 80 sessions. No average differences were found for B_t values, with the exception of sampling errors being higher with lower sample sizes (an intuitive and expected finding). Finally, Equation 2 had no substantial effect on $\log c$, which remained, on average, .00 (Caron, 2014).

Matching behaviors were partially explained without the implementation of any behavioral process nor any effect of reinforcers on behavior. In other words, learning was not necessary to obtain relatively high average explained variance for the GML. The occurrence of responses and reinforcers were independent and highly correlated.

This first part of the study investigates the first environmental constraint; that the amount of reinforcers cannot be higher than the number of responses. The next simulation assesses the influence of Equation 3 on the GML, i.e., the theoretical number of possible responses by sessions (B_t). As pointed earlier, an organism cannot emit an infinite number of responses in a finite condition. Thus, in a given session, the response rate is bounded.

3 Numerical simulation 2

The second Monte Carlo simulation was based on the first. The purpose of the simulation was to assess the effect of implementing Equation 3, i.e., a maximum number of possible responses and reinforcers in a single session, on the parameters of the GML. The simulation was carried out in Matlab (2012a) and was as follows:

- (1) Generate two samples of observed behavior (B_i) with length n from a discrete uniform distribution ($1, B_t$);
- (2) Generate another sample such as $B_i - B_{1j} = B_{2j}$;
- (3) Generate two samples of observed reinforcers from discrete uniform distribution ($1, B_{ij}$), such as $R_{ij} \leq B_{ij}$;
- (4) Fit Equation 1 to the data set and record a , $\log c$ and the r^2 ;
- (5) Repeat steps 1, 2, 3, and 4 a total of 10,000 times.

Like simulation 1, the number of sessions (n) and the maximum bound of the number of responses (B_t) varied from 10, 20, 30, 40, and 50, and 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100, respectively.

3.1 Results

The first simulation revealed an average expected explained variance of .47. Adding the second constraint within the Monte Carlo method in numerical simulation 2 increased the average to .63. In other words, with no behavioral process implemented, the GML explained on average 63% of the variance. There were substantial influences of B_t on the explained variances, but notably the difference is only perceptible at the third decimal. As can be seen in the bottom panel of Figure 3, average slopes were more uniformly ordered; no difference between the B_t values are observable. The

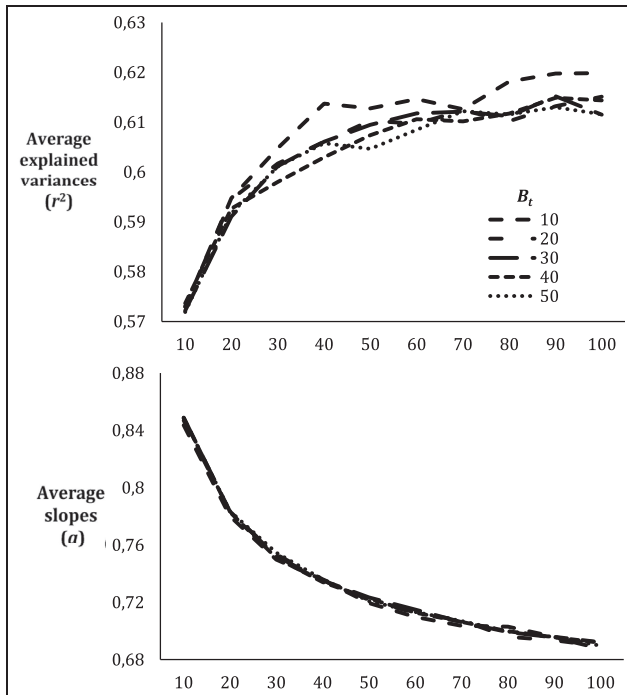


Figure 3. Results from simulation 2. Top panel depicts the average explained variances of the generalized matching law (GML) according to the number of sessions (n) (abscissa) and the upper bound of the uniform distribution (B_t) (line style). The expected explained variance reaches an asymptote approximately at $n=40$, that finally reaches .61 at $n=80$. There were less influences of B_t (compared with Figure 2). Bottom panel presents the average slopes according to the same conditions (abscissa and line style). Average slope decelerated pattern was near perfectly orderly.

average slope in the second simulation is .76 compared with .57 in the first simulation, which is not surprising given the increase in explained variance. Like the numerical simulation 1, average $\log c$ was still .00 for all conditions (varying at the third decimal).

In sum, adding an environmental constraint on the generation of behavior and reinforcers increased the explained variance and the sensitivity (slope) in simulation 2. As in simulation 1, the GML accounted for most of the variance without any programmed learning process. In fact, it is surprising that two simple constraints (Equations 2 and 3) had such an influence on the parameters of the GML.

4 General discussion

The purpose of the current study was to assess the influence of environmental constraints on matching behavior. According to the current findings, concurrent reinforcement schedules clearly induced a quantitative dependency between behaviors and reinforcers. These

simulations show that, as simple as constraints appear, they have a substantial impact on the GML's parameters. The Monte Carlo simulations revealed that the GML explained on average 47% of the variance when Equation 2 was implemented, and increased to 63% when Equation 3 was added. Environmental constraints likely have an influence within feedback systems: an effect substantial enough to generate 63% of explained variance by the GML. Matching behavior can thus be conceptualized not only as the result of a behavioral (decision-making) process, but as the product of environmental conditions.

This paper sought to investigate whether computational learning models are necessary to produce matching behavior. The answer is partly yes. Compared with the results of the current computation of environmental constraints, the GML reaches much better explained variances with real (and artificial) organisms especially in experimental studies. In most experiments with animal and human subjects, the GML accounts for a large proportion of variance (higher than .80) and the parameter a varies around .80 (Baum, 1979; Davison & McCarthy, 1988; McDowell, 2013a). In the second numerical simulation, in which the number of responses was constrained, the GML accounted for .63 of the variance and the parameter a varied around .76 on average. The average slope, the parameter a , was approximately the same as for biological agents (.76), whereas the average explained variance was lower (.63). The gap between simulations and experimental studies point out that, at least, a decision-making process would be necessary to explain all the variance accounted by the GML. In real-life settings, the variance cannot be distinguished between environmental constraints and behavioral processes. Both of these could account for matching behavior. At the very least, these simulations show that environmental constraints can spuriously increase the explained variance accounted by the GML. Their influence in real-life settings has yet to be experimentally tested, regardless of the behavioral process taking place.

The current results do not enlighten the current model of living-agents matching behavior. However, simulations point out some of the limitations in the actual view of matching behavior. First, they show that matching behavior is not merely an emergent property of behavioral processes, but can be an artefact shaped by environmental constraints. At best, they suggest a new computational perspective of matching behavior: modeling the environment rather than the organism. This leads to a Skinnerian account of the matching phenomenon: the content of the black box will not be the main point of interest, but rather the environment that acts upon it. The standpoint advocated by Davison and McCarthy (1988) and Herrnstein (1997) that any behavioral model of choice should first account for

matching behavior (such as melioration theory and other related models) is weakened. Results challenge matching behavior per se and strengthen an interpretation of matching behavior as a product of a feedback system. This new direction can and should be investigated.

Finally, the current findings can be particularly useful for data gathered for real-life descriptive analyses of organisms (i.e., observing organisms' behaviors with only natural variations and without any experimental control), in which the likelihood of the operant function of reinforcers (and consequently matching behaviors) is unknown (Caron, 2014). In natural settings, it is impossible to conclude without doubt whether, for example, a child's behavior is at a steady state, whether its behavior are reinforced by the contingencies and what behavioral processes are taking place (Bijou, Peterson & Holt, 1968). The current findings, that the expected average variance explained by the GML is .63 with a slope of .76 on average, could be used as a lower bound or a null hypothesis to assess whether the GML accounts for the data better than a pure random process. For instance, Rivard, Forget, Kerr and Bégin (2014) analyzed the matching behavior of 14 children with autism spectrum disorder in order to assess their sensitivity to attention from their therapist. They used a rule of thumb to evaluate whether children's behaviors were sensitive. If the GML explained more than .50 of the variance of children's behaviors (Reed, 2009), they were concluded to be sensitive. In this example, using the current simulation results would lead to the use of an empirically orientated test rather than a rule of thumb. Though their conclusions in this instance would remain relatively unaffected, their results would be more accurate and precise, and the confidence in them would increase.

In summary, computational learning models appear in part necessary to produce matching behavior. Based on these findings, there is evidence that other environmental constraints could increase the average explained variance accounted for by the GML. The current simulations investigated two possible environmental constraints, though others can be conceptualized. Future simulations could explore if constraints restraining choice lead to more robust outcomes or higher explained variances. A more comprehensive and exhaustive description of environmental constraints is needed to account for matching behaviors, whether it concerns concurrent schedules of reinforcement or a two-armed bandit. Still, it opens more avenues toward an environmentalism-orientated behaviorism rather than a cryptocognivism.

Acknowledgments

I thank Philippe Valois, Jacques Forget, and Stephanie Caldas for their comments on earlier drafts.

Conflict of interest

Author reports no conflict of interest.

Funding

The author is currently subsidized by the Fonds de recherche du Québec—Société et culture.

Notes

1. Herein, the word "constraint" designates "restriction of variability". It may be understood as "systematic bias", "limitation", or "boundary".
2. Notice that the study was published in French.
3. Remember that the GML is a within-subject model (Caron, 2013) and that the sample size is the number of sessions by subject.

References

- Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, 20, 137–153.
- Baum, W. M. (1974). On two types of deviation from the matching law: Bias and undermatching. *Journal of the Experimental Analysis of Behavior*, 22, 231–242.
- Baum, W. M. (1979). Matching, undermatching, and overmatching in studies of choice. *Journal of the Experimental Analysis of Behavior*, 32, 269–281.
- Bijou, S. W., Peterson, R. F., & Ault, M. H. (1968). A method to integrate descriptive and experimental field studies at the level of data and empirical concepts. *Journal of Applied Behavior Analysis*, 1, 175–191.
- Caron, P.-O. (2013). On applying the matching law to between-subject data. *Animal Behaviour*, 85, 857–860.
- Caron, P.-O. (2014). La loi généralisée de l'appariement: une simulation de Monte-Carlo. *Acta Comportamentalia*, 22, 169–179.
- Davison, M., & McCarthy, D. (1988). *The matching law: A research review*. Hilldale, NJ: Erlbaum.
- Dawson, M. R. W., Dupuis, B., Spetch, M. L., & Kelly, D. M. (2009). Simple artificial neural networks that match probability and exploit and explore when confronting a multiarmed bandit. *IEEE Transactions on Neural Networks*, 20, 1368–1371.
- Gallistel, C. R. (2005). Deconstructing the law of effect. *Games and Economic Behavior*, 52, 410–423.
- Herrnstein, R. J. (1997). *The matching law. Papers in psychology and economics*. Cambridge, MA: Harvard University Press.
- Iigaya, K., & Fusi, S. (2013). Dynamical regimes in neural network models of matching behavior. *Neural Computation*, 25, 3093–3112.
- Loewenstein, Y. (2008). Robustness of learning that is based on covariance-driven synaptic plasticity. *PLoS Computational Biology*, 4, e1000007.
- McDowell, J. J. (2004). A computational model of selection by consequences. *Journal of the Experimental Analysis of Behavior*, 81, 297–317.
- McDowell, J. J. (2013a). On the theoretical and empirical status of the matching law and matching theory. *Psychological Bulletin*, 139, 1000–1028.

- McDowell, J. J. (2013b). A quantitative evolutionary theory of adaptive behavior dynamics. *Psychological Review*, *120*, 731–750.
- McDowell, J. J., & Ansari, Z. (2005). The quantitative law of effect is a robust emergent property of an evolutionary algorithm for reinforcement learning. In M. Caparrere, A. Freitas, P. Bentley, C. Johnson, & J. Timmis (Eds.), *Advances in artificial life: ECAL 2005, LNAI 3630*. (pp. 413–422). Berlin, Germany: Springer-Verlag.
- Rachlin, H., Battalio, R. C., Kagel, J. H., & Green, L. (1981). Maximization theory in behavioral psychology. *The Behavioral and Brain Sciences*, *4*, 371–388.
- Reed, D. D. (2009). Using Microsoft Office Excel® 2007 to conduct generalized matching analyses. *Journal of Applied Behavior Analysis*, *42*, 867–875.
- Rivard, M., Forget, J., Kerr, K., & Bégin, J. (2014). Matching law and sensitivity to therapist's attention in children with autism spectrum disorders. *The Psychological Record*, *65*, 79–88.
- Saito, H., Katahira, K., Okanoya, K., & Okada, M. (2014). Bayesian deterministic decision making: A normative account of the operant matching law and heavy-tailed reward history dependency of choices. *Frontiers in Computational Neuroscience*, *8*, 1–7.
- Sakai, Y., & Fukai, T. (2008). When does reward maximization lead to matching law? *PLoS One*, *3*, e3795.
- Skinner, B. F. (1938). *The behavior of organism*. New York: Appleton Century Croft.
- Shimp, C. P. (1992). Computational behavior dynamics: An alternative description of Nevin (1969). *Journal of the Experimental Analysis of Behavior*, *57*, 289–299.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.
- Wearden, J. H., & Burgess, I. S. (1982). Matching since Baum (1979). *Journal of the Experimental Analysis of Behavior*, *38*, 339–348.

About the Author



Pier-Olivier Caron is a PhD student in psychology and is currently supervised by Jacques Forget DPs. His main interests are in quantitative models from the experimental analysis of behavior, multivariate data analyses, statistical modeling, and computation. He is currently working on the generalized matching law and its applications for the assessment of social sensitivity.