Check for updates

Fish and chips: Using machine learning to estimate the effects of basal cortisol on fish foraging behavior

Wallace M. Bessa[1]*, Lucas S. Cadengue[2] and Ana C. Luchiari[3]

[1]Department of Mechanical and Materials Engineering, University of Turku, Turku, Finland, [2]Programa de Pós-Graduação em Engenharia Elétrica e de Computação, Universidade Federal do Rio Grande do Norte, Natal, Brazil, [3]Departamento de Fisiologia, Centro de Biociências, Universidade Federal do Rio Grande do Norte, Natal, Brazil

Foraging is an essential behavior for animal survival and requires both learning and decision-making skills. However, despite its relevance and ubiquity, there is still no effective mathematical framework to adequately estimate foraging performance that also takes interindividual variability into account. In this work, foraging performance is evaluated in the context of multi-armed bandit (MAB) problems by means of a biological model and a machine learning algorithm. Siamese fighting fish (*Betta splendens*) were used as a biological model and their ability to forage was assessed in a four-arm cross-maze over 21 trials. It was observed that fish performance varies according to their basal cortisol levels, i.e., a reduced average reward is associated with low and high levels of basal cortisol, while the optimal level maximizes foraging performance. In addition, we suggest the adoption of the epsilon-greedy algorithm to deal with the exploration-exploitation tradeoff and simulate foraging decisions. The algorithm provided results closely related to the biological model and allowed the normalized basal cortisol levels to be correlated with a corresponding tuning parameter. The obtained results indicate that machine learning, by helping to shed light on the intrinsic relationships between physiological parameters and animal behavior, can be a powerful tool for studying animal cognition and behavioral sciences.

KEYWORDS

foraging, fish, multi-armed bandit, epsilon-greedy, cortisol

1. Introduction

Foraging plays an essential role in animal's fitness and its efficacy has been shaped by evolutionary processes (Pearson et al., 2014). However, in addition to its phylogenetic basis, foraging behavior is also largely influenced by ontogeny (Hughes et al., 1992; Grecian et al., 2018). In fact, both intrinsic and extrinsic factors that modulate behavior confer different consequences on fitness (Dingemanse et al., 2004; Brown et al., 2007) and the huge variation between individuals' life history leads to several differences in the way they deal with stressful situations, how sensitive to changes they are, or even how they learn new tasks.

But how do physiological responses correlate with these factors and their corresponding behavioral patterns? In this context, Koolhaas et al. (1999) defined two main profiles: proactive individuals show high exploration, are less sensitive to environmental changes, have low latency to aggression and low hypothalamic-pituitary-adrenal (HPA) response, while reactive are less explorative, highly dependent on environmental stability, have less routine development and high HPA reactivity. Based on an extensive literature review, Toscano et al. (2016) suggested that these behavioral traits and their interindividual variations can lead to specialization in foraging

behavior. Cortisol levels are indeed highly influential on many traits, including immune response (Zhang et al., 2022), reproductive investment (Jiang et al., 2022), metabolism (Fernandes Silva et al., 2022; Winberg and Sneddon, 2022), energy reallocation to cope with stressors (Gorissen and Flik, 2016), and the behavioral and cognitive responses (Alfonso et al., 2019, 2020; Torgerson-White and Sánchez-Suárez, 2022). Furthermore, it is also important to emphasize that in many animal species, basal and post-stress cortisol levels, as well as behavioral responses, are correlated in such a way that these characteristics are used to identify coping styles (Alfonso et al., 2020).

Recently, Morimoto (2019) showed that foraging decisions can be represented in the context of multi-armed bandit (MAB) problems. The MAB problem is a machine learning framework used to investigate the exploration vs. exploitation dilemma: the learning agent should try to maximize the cumulative reward by choosing between exploitation (an immediate payoff based on available information) or exploration (looking for options that may lead to a better payoff). Reinforcement learning algorithms proposed to deal with this problem usually rely on a parameter to modulate the agent exploitation/exploitation tradeoff (Sutton and Barto, 2018). Moreover, it is already known that different species have developed different strategies to balance exploration and exploitation depending on the ecological conditions of their environment. Within the same species, these strategies may be modulated by specific internal states and their corresponding physiological parameters. Therefore, assuming that interindividual physiological differences lead to distinct foraging behaviors, would it be possible to estimate foraging performance based on physiological indicators?

In this work, we propose that basal cortisol can be used as a predictor of the foraging behavior of the Siamese Fighting Fish, *Betta splendens*, as well as its performance in MAB related tasks. Despite their ancient evolutionary appearance, fishes stand for good biological models. Their brain structure and rich behavioral repertoire are comparable to those of mammals (Gerlai, 2014), which offers many possibilities for research on vertebrate evolution, including both cognitive and behavioral comparative analyses. To model foraging decisions, we apply the ε-greedy algorithm (Sutton and Barto, 2018) and suggest that the parameter ε, used to modulate the exploration/exploitation tradeoff, correlates with normalized basal cortisol levels of fish. The close agreement between the results obtained with biological and computational models also allows us to suggest that simple algorithms like this one can not only describe quite complex behavioral patterns, but also characterize the decision mechanisms of many organisms.

2. Materials and methods

2.1. Biological model

Eighty-six male *Betta splendens* (adults, 3 months old, acquired in a local breeding facility in Natal/RN) were kept in individual tanks (3 L, $15 \times 15 \times 15$ cm^3) in the vivarium of the Fish Laboratory at the Federal University of Rio Grande do Norte throughout the experimental phase. They were weighed and measured (0.87 ± 0.15 g and 2.70 ± 0.21 cm) before being placed in the tanks. The tanks were arranged side by side with an opaque partition between them to avoid visual contact between the fish. The water in the tanks was changed every 2 days and its quality was monitored daily: $28 \pm 1^{\circ}$C,

TABLE 1 Fish groups divided by basal cortisol level.

Group	Sample size	Cortisol range (in ng/mL)	Normalized cortisol (mean ± SE)
1	16	1.68–10.69	0.0289 ± 5.22E-03
2	17	12.78–28.55	0.1413 ± 9.85E-03
3	16	32.18–87,35	0.4103 ± 4.32E-02
4	16	90.50–117.33	0.8137 ± 1.53E-02
5	17	117.44–135.40	0.9251 ± 9.74E-03

pH ~ 6.7 and O$_2$ $\sim$ 6 mg/L. The photoperiod was set on a 12:12 light/dark cycle, and the animals were fed *ad libitum* twice a day with commercial pellet food (Alcon Betta, 44% protein and 5% fat) and frozen *Artemia salina*.

In the foraging task, a cross-maze was used: 10 cm high, with a central area of 10×10 cm^2, and four arms measuring 30 cm long by 10 cm wide each. Each fish was individually placed in the central area, in a lift-up start box, for 60 s and then released to explore the maze. One of the cross-maze arms had a small red circular plate (3 cm diameter) on which five units of *Artemia salina* were placed when fish entered the area (reward area). The reward arm was randomly chosen for each trial to prevent the fish from learning any other cues other than the red circular plate. This procedure was performed 21 times (trials) over 11 days. A reward equal to one was assigned to each trial in which the fish entered first the arm with the circular plate (success), and equal to zero in the other cases (failure). There were no cues on the maze walls but the feeding plate. A black curtain was used to hide the researcher from the fish during the trials. After each trial, the fish was removed from the maze and returned to its original tank. All trials were recorded using a camcorder (JVC Full HD 60x, model GZ-HM65BU) located above the tank.

To estimate basal cortisol levels, the fish were euthanized 24 h after the last trial, all at the same time, with a high dose of anesthetic (clove oil, 10 mL/L) and stored in individual Falcon tubes at -20°C. Each whole-body sample was macerated with 3 mL of phosphate-buffered saline (PBS) using a high-speed stirrer (Nova Tecnica Homogenizador Potter NT 136). The resulting solution was then centrifuged (e.g., for 10 min at 3000 g and 7°C). To standardize fish samples, 1 dL aliquots of supernatant from each fish homogenate were collected and transferred to Eppendorf tubes. A salivary cortisol ELISA kit (SLV 2930 Lot 64K044 DRG Diagnostics) was used. Both cortisol extraction and kit validation followed Sink et al. (2007). The cortisol test failed for three individuals and for another one it presented an outlier, resulting in 82 valid measurements. Moreover, to facilitate comparative analysis, the cortisol levels were rescaled to range from zero to one using min-max normalization and then the 82 fish were divided into five groups (2 with 17 individuals, 3 with 16 individuals). Thus, in order to allow a better visualization of the obtained results and facilitate the identification of different foraging patterns, the fish were divided into two groups with low levels of basal cortisol, two with high levels and an intermediate group. Table 1 shows the fish groups with their corresponding sample sizes, measured cortisol ranges (in ng/mL) and normalized values (group means with standard errors). An OpenDocument Spreadsheet with the foraging data is available as Supplementary material.

2.2. Computational model

The ε-greedy algorithm relies on a single parameter ε to define the exploratory bias. The parameter can range from 0 to 1, which directly corresponds to a continuum from high exploitation to high exploration. At each iteration (trial), the algorithm chooses between exploiting (with probability $1 - \varepsilon$) the best arm so far or exploring (with probability ε) a new one. After choosing an arm A, the algorithm assigns the corresponding reward R to this action, updates the number of times this arm was visited $N(A)$ and its action-value $Q(A)$. The value $Q(A)$ of an arm A represents the mean reward for this action up to the current trial. After visiting the arm in a trial and getting the corresponding reward, the action-value is updated as follows Sutton and Barto (2018):

$$Q(A) \leftarrow Q(A) + \frac{1}{N(A)}\Big[R - Q(A)\Big]. \tag{1}$$

The ε-greedy algorithm for a bandit with k arms is presented in the box below and its implementation in Python is available as Supplementary material. In the simulation studies, to replicate the foraging task, four arms were considered $a = [1, 2, 3, 4]$ but only one of them yielded a reward $R = 1$ (for the other three arms there was no payoff, $R = 0$). All fish had their computer versions properly run, with each epsilon being assigned the corresponding normalized cortisol level. Two sets of simulations were performed for the 82 epsilons: first with 21 trials for each ε (for comparative analysis with fish) and then with 400 trials (to assess their long-term performance).

```
Set exploration probability ε
Initialize N(a) ← 0, for a from 1 to k
Initialize Q(a) ← 0, for a from 1 to k
repeat
    Pick a random number p between 0 and 1
    if p < ε then
        Choose an arm at random:
        A ← random(a)

    else
        Choose best arm so far:
        A ← arg max_a Q(a)

    end
    Get reward of the chosen arm:
    R ← reward(A)
    Update N(A) ← N(A) + 1
    Update Q(A) ← Q(A) + [R − Q(A)]/N(A)

until last trial
```

Algorithm 1. ε-greedy for a k-armed bandit

3. Results and discussion

The foraging performance of both biological and computational models after 21 trials is shown in Figure 1 and Table 2. The data (mean $\pm$ SE) are presented for each group based on the average rewards obtained by the individuals. For each individual, the average reward stands for the sum of its obtained payoffs divided by the maximum possible score (21 in this case).
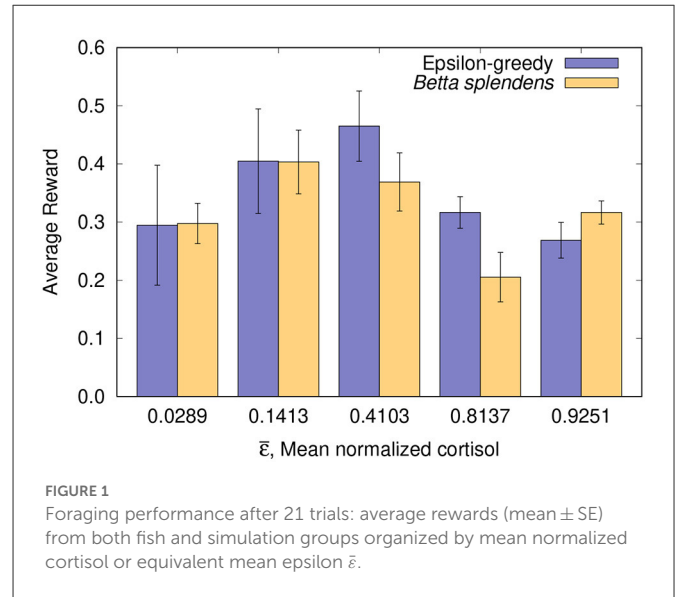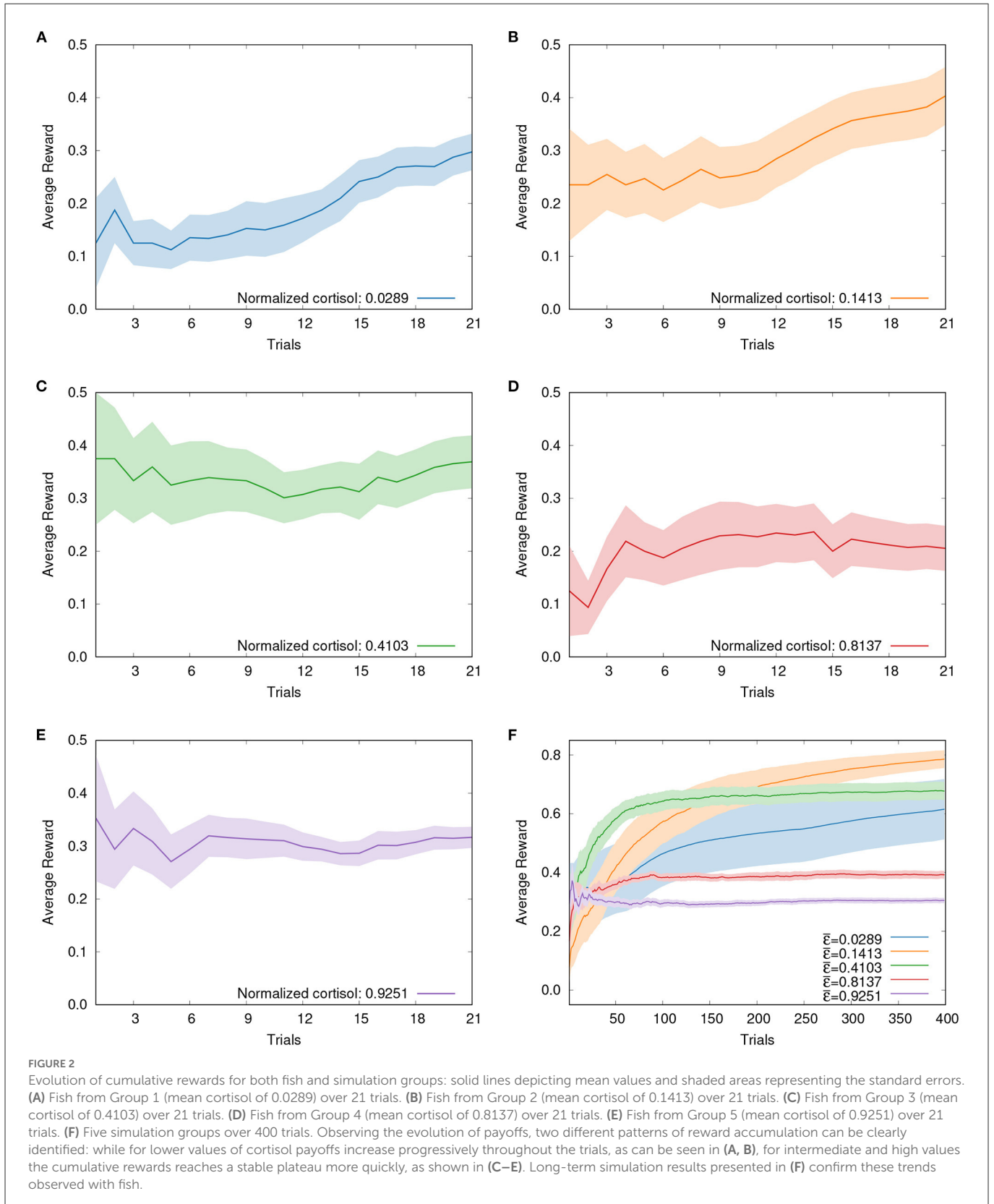

FIGURE 1
Foraging performance after 21 trials: average rewards (mean $\pm$ SE) from both fish and simulation groups organized by mean normalized cortisol or equivalent mean epsilon $\bar{\varepsilon}$.

TABLE 2 Average rewards from both fish and simulation groups after 21 trials.

Group	ε-greedy (mean $\pm$ SE)	*Betta splendens* (mean $\pm$ SE)
1	0.2946 $\pm$ 0.1031	0.2976 $\pm$ 0.0346
2	0.4048 $\pm$ 0.0899	0.4034 $\pm$ 0.0548
3	0.4650 $\pm$ 0.0605	0.3690 $\pm$ 0.0500
4	0.3165 $\pm$ 0.0271	0.2054 $\pm$ 0.0425
5	0.2687 $\pm$ 0.0307	0.3165 $\pm$ 0.0200

It can be seen in Figure 1 and Table 2 that the results obtained with the fish and with the corresponding computer simulations are in close agreement, which allows us to suggest that the ε-greedy algorithm is indeed capable of estimating the foraging performance of the Siamese Fighting Fish, *Betta splendens*. Moreover, it also supports the assumption that the algorithm can be used to investigate the effects of physiological parameters on foraging behavior. By observing the results in Figure 1, it can be verified that the normalized cortisol level and the parameter ε show a good correlation with respect to the average reward obtained in the corresponding groups. In both cases, a reduced average reward is associated with low and high values of ε (or the normalized basal cortisol, in the biological model), while an optimal value maximizes foraging performance. In fact, this variability can be explained from the algorithm's point of view, since ε is linked to exploratory bias, and its optimal value is related to the best exploration/exploitation tradeoff for each foraging task.

As expected, extreme levels of basal cortisol, both on the right and on the left, lead to a lower foraging performance, according to an inverted-U pattern. This inverted-U shape is commonly seen in the relationship between glucocorticoids and performance in cognitive functions such as learning and decision making (Salehi et al., 2010; Schilling et al., 2013). As a matter of fact, both low and high cortisol levels are linked to cognitive dysfunction (Maripuu et al., 2014). It has been shown that cases of major depression correlate with a very low level of cortisol (Keilp et al., 2016), which cannot stimulate

FIGURE 2
Evolution of cumulative rewards for both fish and simulation groups: solid lines depicting mean values and shaded areas representing the standard errors.
(A) Fish from Group 1 (mean cortisol of 0.0289) over 21 trials. **(B)** Fish from Group 2 (mean cortisol of 0.1413) over 21 trials. **(C)** Fish from Group 3 (mean cortisol of 0.4103) over 21 trials. **(D)** Fish from Group 4 (mean cortisol of 0.8137) over 21 trials. **(E)** Fish from Group 5 (mean cortisol of 0.9251) over 21 trials. **(F)** Five simulation groups over 400 trials. Observing the evolution of payoffs, two different patterns of reward accumulation can be clearly identified: while for lower values of cortisol payoffs increase progressively throughout the trials, as can be seen in **(A, B)**, for intermediate and high values the cumulative rewards reaches a stable plateau more quickly, as shown in **(C−E)**. Long-term simulation results presented in **(F)** confirm these trends observed with fish.

normal body and brain functioning, including motivation. On the other hand, high cortisol levels are related to chronic stress, and also associated with poor performance (Lupien et al., 2018). Thus, inspecting the cortisol data, such variation was already expected, as

well as its relationship with the performance of the fish in the learning task. The biological relevance here concerns the ability of animals to learn and respond appropriately to environmental demands, i.e., animals that are in a situation of low or high levels of cortisol (derived

from some previous situation experienced by them) may present performance impairment and threat to fitness.

Like cortisol, other factors that influence foraging can affect the animal's ability or the way it learns, for example, the speed to make associations (Wang et al., 2015). As suggested by Raoult et al. (2017), low-stress fish show higher levels of activity and explore the environment more than high-stress fish, which leads to learning the reward location as a result of their proactive coping style (Koolhaas et al., 1999).

It should be noted that the slight discrepancies observed for intermediate and higher values of mean epsilon/cortisol in Figure 1 and Table 2 are likely due to the low number of trials, which can lead to a greater performance variability, especially in cases where a more exploratory behavior is involved. Moreover, in Group 4 (with corresponding mean normalized cortisol of 0.8137) it happened that 25% of the individuals (4 out of 16) did not accumulate any reward, which ended up reducing the average reward of the entire group and increasing its discrepancy in relation to computer estimate. For comparison purposes, in the other groups, at most one fish did not accumulate a reward, which led to a closer agreement between the data obtained with the fish and the algorithm. In order to make this relationship even clearer, the correlation coefficient between the data obtained with the fish and with the epsilon-greedy algorithm can be calculated for two different scenarios:

1. With all groups included, where $r = 0.62$ (moderately correlated) is obtained;
2. With all groups except Group 4, where $r = 0.81$ (highly correlated) is obtained.

As a matter of fact, it is known that foraging performance can be improved through learning, which in fish is most easily measured by the number of exposures required (Warburton, 2003). However, this learning efficiency is heavily influenced by the complexity of the task, which can make the minimum number of trials range from <10 to more than 200 to reach an asymptote (Hughes et al., 1992). For practical reasons, associative tasks involving fish models are usually repeated 20 times (Luchiari et al., 2015) of which significant behavioral changes (learning) begin to occur around the 12th trial. In addition, internal factors such as hormone levels also affect the animals' performance.

Furthermore, if the evolution of cumulative rewards over trials is inspected, two distinct patterns can be clearly observed. The results (mean $\pm$ SE) obtained with *Betta splendens* and ε-greedy are presented in Figure 2. In Figures 2A, B, for instance, it can be seen that for lower values of basal cortisol the payoffs increase progressively over the trials, while for intermediate and high values the cumulative rewards reaches a stable plateau more quickly, as shown in Figures 2C–E. The ε-greedy algorithm in this context can also help to confirm this trend in the long term (see Figure 2F). The simulation of the five groups along 400 trials shows that the best foraging performance is shifted to the left as the number of trials increases. That is, if we select specific numbers of trials, we will see that the best foraging performance, as the number of trials increases, will gradually be obtained by individuals with lower levels of basal cortisol. This suggests that the exploratory behavior associated with higher basal cortisol levels is initially favored, but as the number of trials increases, individuals with lower basal cortisol levels, i.e., prone to exploitation, tend to perform better. Further

experiments involving more trials, different types of mazes and even other biological models are expected to corroborate this conjecture.

It is worth mentioning that our biological results echo work on individuals' differences in speed-accuracy tradeoffs (Raoult et al., 2012, 2017; Mamuneas et al., 2015), indicating that corticosteroids modulate the learning performance. It also suggests that individuals with different basal corticoid levels employ distinct learning strategies and may need different number of trials to reach the same level of performance.

Moreover, our findings endorse that multi-armed bandits can be a reasonable framework to evaluate foraging decisions, as suggested by Morimoto (2019). We show that the ε-greedy turns out to be a simple tool but capable of describing both qualitatively and quantitatively the animal behavior in foraging tasks. By means of a stochastic approach, the algorithm allows the exploratory behavior to be uniformly spread across all trials. This is an essential feature to capture the ability of animals to make decisions in an ever-changing dynamic environment.

Finally, it is important to highlight that this synergy between computer and biological sciences can benefit both areas. In fact, much of the work done on machine learning has been guided by the research on animal cognition and behavior. Biologically inspired schemes have indeed allowed the development of more efficient algorithms for autonomous robots (Bessa et al., 2017, 2018). Furthermore, considering that decision-making algorithms are already quite widespread in our society, with applications ranging from targeted advertisements and recruitment choices to medical diagnoses and prison sentences, it is crucial and urgent to better understand their capabilities and limitations. In this sense, animal behavior can provide powerful insights and end up helping to design more efficient and ethical algorithms.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The animal study was reviewed and approved by Ethics Committee for Animal Use of the Federal University of Rio Grande do Norte (CEUA-006/2016).

Author contributions

WB, LC, and AL contributed to conception and design of the study. AL organized the fish database. LC implemented the computer code and performed the computational analysis. WB and AL wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

(CAPES) and the Brazilian National Council for Scientific and Technological Development (CNPq).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnbeh.2023.1028190/full#supplementary-material

The OpenDocument Spreadsheet with the foraging data and the Python implementation of the ε-greedy algorithm for a bandit with k arms are available as Supplementary material.

References

Alfonso, S., Sadoul, B., Gesto, M., Joassard, L., Chatain, B., Geffroy, B., et al. (2019). Coping styles in european sea bass: the link between boldness, stress response and neurogenesis. *Physiol. Behav.* 207, 76–85. doi: 10.1016/j.physbeh.2019.04.020

Alfonso, S., Zupa, W., Manfrin, A., Fiocchi, E., Spedicato, M. T., Lembo, G., et al. (2020). Stress coping styles: Is the basal level of stress physiological indicators linked to behaviour of sea bream? *Appl. Anim. Behav. Sci.* 231, 105085. doi: 10.1016/j.applanim.2020.105085

Bessa, W. M., Brinkmann, G., Duecker, D. A., Kreuzer, E., and Solowjow, E. (2018). A biologically inspired framework for the intelligent control of mechatronic systems and its application to a micro diving agent. *Math. Probl. Eng.* 2018, 9648126. doi: 10.1155/2018/9648126

Bessa, W. M., Kreuzer, E., Lange, J., Pick, M.-A., and Solowjow, E. (2017). Design and adaptive depth control of a micro diving agent. *IEEE Robot. Autom. Lett.* 2, 1871–1877. doi: 10.1109/LRA.2017.2714142

Brown, C., Burgess, F., and Braithwaite, V. A. (2007). Heritable and experiential effects on boldness in a tropical poeciliid. *Behav. Ecol. Sociobiol.* 62, 237–243. doi: 10.1007/s00265-007-0458-3

Dingemanse, N. J., Both, C., Drent, P. J., and Tinbergen, J. M. (2004). Fitness consequences of avian personalities in a fluctuating environment. *Proc. R. Soc. B: Biol. Sci.* 271, 847–852. doi: 10.1098/rspb.2004.2680

Fernandes Silva, B. W., Leite-Ferreira, M. E., Menezes, F. P., and Luchiari, A. C. (2022). Covariation among behavioral traits and hatching time in zebrafish. *Behav. Process.* 194, 104546. doi: 10.1016/j.beproc.2021.104546

Gerlai, R. (2014). Fish in behavior research: Unique tools with a great promise! *J. Neurosci. Methods* 234, 54–58. doi: 10.1016/j.jneumeth.2014.04.015

Gorissen, M., and Flik, G. (2016). "The endocrinology of the stress response in fish: an adaptation-physiological view," in *Biology of Stress in Fish, volume 35 of Fish Physiology*, eds C. B. Schreck, L. Tort, A. P. Farrell, and C. J. Brauner (Cambridge, MA: Academic Press), 75–111.

Grecian, W. J., Lane, J. V., Michelot, T., Wade, H. M., and Hamer, K. C. (2018). Understanding the ontogeny of foraging behaviour: insights from combining marine predator bio-logging with satellite-derived oceanography in hidden markov models. *J. R. Soc. Interface* 15, 20180084. doi: 10.1098/rsif.2018.0084

Hughes, R. N., Kaiser, M. J., Mackney, P. A., and Warburton, K. (1992). Optimizing foraging behaviour through learning. *J. Fish Biol.* 41, 77–91. doi: 10.1111/j.1095-8649.1992.tb03870.x

Jiang, P., Guo, J., Chen, A., and Lu, W. (2022). Reproductive endocrine variation in female olive flounder with different stress coping styles during the breeding season. *Aquac. Res.* 53, 3949–3961. doi: 10.1111/are.15898

Keilp, J. G., Stanley, B. H., Beers, S. R., Melhem, N. M., Burke, A. K., Cooper, T. B., et al. (2016). Further evidence of low baseline cortisol levels in suicide attempters. *J. Affect. Disord.* 190, 187–192. doi: 10.1016/j.jad.2015.10.012

Koolhaas, J., Korte, S., De Boer, S., Van Der Vegt, B., Van Reenen, C., Hopster, H., et al. (1999). Coping styles in animals: current status in behavior and stress-physiology. *Neurosci. Biobehav. Rev.* 23, 925–935. doi: 10.1016/S0149-7634(99)00026-3

Luchiari, A. C., Salajan, D. C., and Gerlai, R. (2015). Acute and chronic alcohol administration: Effects on performance of zebrafish in a latent learning task. *Behav. Brain Res.* 282, 76–83. doi: 10.1016/j.bbr.2014.12.013

Lupien, S. J., Juster, R.-P., Raymond, C., and Marin, M.-F. (2018). The effects of chronic stress on the human brain: From neurotoxicity, to vulnerability, to opportunity. *Front. Neuroendocrinol.* 49, 91–105. doi: 10.1016/j.yfrne.2018.02.001

Mamuneas, D., Spence, A. J., Manica, A., and King, A. J. (2015). Bolder stickleback fish make faster decisions, but they are not less accurate. *Behav. Ecol.* 26, 91–96. doi: 10.1093/beheco/aru160

Maripuu, M., Wikgren, M., Karling, P., Adolfsson, R., and Norrback, K.-F. (2014). Relative hypo- and hypercortisolism are both associated with depression and lower quality of life in bipolar disorder: a cross-sectional study. *PLoS ONE* 9, e98682. doi: 10.1371/journal.pone.0098682

Morimoto, J. (2019). Foraging decisions as multi-armed bandit problems: applying reinforcement learning algorithms to foraging data. *J. Theor. Biol.* 467, 48–56. doi: 10.1016/j.jtbi.2019.02.002

Pearson, J., Watson, K., and Platt, M. (2014). Decision making: the neuroethological turn. *Neuron* 82, 950–965. doi: 10.1016/j.neuron.2014.04.037

Raoult, V., Brown, C., Zuberi, A., and Williamson, J. E. (2012). Blood cortisol concentrations predict boldness in juvenile mulloway (argyosomus japonicus). *J. Ethol.* 30, 225–232. doi: 10.1007/s10164-011-0314-9

Raoult, V., Trompf, L., Williamson, J. E., and Brown, C. (2017). Stress profile influences learning approach in a marine fish. *PeerJ.* 5, e3445. doi: 10.7717/peerj.3445

Salehi, B., Cordero, M. I., and Sandi, C. (2010). Learning under stress: the inverted-u-shape function revisited. *Learn. Mem.* 17, 522–530. doi: 10.1101/lm.1914110

Schilling, T. M., Kölsch, M., Larra, M. F., Zech, C. M., Blumenthal, T. D., Frings, C., et al. (2013). For whom the bell (curve) tolls: cortisol rapidly affects memory retrieval by an inverted u-shaped dose-response relationship. *Psychoneuroendocrinology* 38, 1565–1572. doi: 10.1016/j.psyneuen.2013.01.001

Sink, T. D., Kumaran, S., and Lochmann, R. T. (2007). Development of a whole-body cortisol extraction procedure for determination of stress in golden shiners, notemigonus crysoleucas. *Fish Physiol. Biochem.* 33, 189–193. doi: 10.1007/s10695-007-9130-0

Sutton, R. S., and Barto, A. G. (2018). *Reinforcement Learning: An Introduction, 2nd Edn*. Cambridge, MA: MIT Press.

Torgerson-White, L., and Sánchez-Suárez, W. (2022). Looking beyond the shoal: Fish welfare as an individual attribute. *Animals* 12, 2592. doi: 10.3390/ani12192592

Toscano, B. J., Gownaris, N. J., Heerhartz, S. M., and Monaco, C. J. (2016). Personality, foraging behavior and specialization: integrating behavioral and food web ecology at the individual level. *Oecologia* 182, 55–69. doi: 10.1007/s00442-016-3648-8

Wang, M.-Y., Brennan, C. H., Lachlan, R. F., and Chittka, L. (2015). Speed-accuracy trade-offs and individually consistent decision making by individuals and dyads of zebrafish in a colour discrimination task. *Anim. Behav.* 103, 277–283. doi: 10.1016/j.anbehav.2015.01.022

Warburton, K. (2003). Learning of foraging skills by fish. *Fish Fish.* 4, 203–215. doi: 10.1046/j.1467-2979.2003.00125.x

Winberg, S., and Sneddon, L. (2022). Impact of intraspecific variation in teleost fishes: aggression, dominance status and stress physiology. *J. Exp. Biol.* 225, jeb169250. doi: 10.1242/jeb.169250

Zhang, X., Yuan, M., and Lu, W. (2022). Individuality in flounder (*Paralichthys olivaceus*): differential responses of immune-endocrine during temperature stress in head kidney. *Aquac. Res.* 53, 1327–1336. doi: 10.1111/are.15667