Please paste or attach your review below.

- 1. Be supportive: Avoid rude, emotive, and accusatory language and tone.
- 2. **Be precise:** In references to manuscript text, analyses and figures, include the line number and figure panels in question.
- 3. Fully reference all of your non-trivial scientific claims (as you would in a manuscript).
- 4. **Review the** *claims*: Perceived novelty or impact of the work must not factor into reviewer comments, unless these form part of the specific claims of the manuscript (e.g., a claim of novelty).
- 5. **Be constructive:** Whenever possible, suggest a solution for the problem or concern you raise. Note, this will always only be one of different possible solutions the authors may choose a different one.

In addition, we encourage you to **answer the following questions** to help recommenders make well-informed and efficient decisions.

If you answer "No" to a question, please **explain why and list your suggestions for improvement** by the authors **in your free-text review or below each question.**

- Title and abstract
 - Does the title clearly reflect the content of the article? [X] Yes, [] No (please explain), [] I don't know
 - Does the abstract present the main findings of the study? [X] Yes, [] No (please explain), [] I don't know

• Introduction

- Are the research questions/hypotheses/predictions clearly presented? [X] Yes, []
 No (please explain), [] I don't know
- Does the introduction build on relevant research in the field? [X] Yes, [] No (please explain), [] I don't know

• Materials and methods

- Are the methods and analyses sufficiently detailed to allow replication by other researchers? [X] Yes, [] No (please explain), [] I don't know
- Are the methods and statistical analyses appropriate and well described? [X] Yes, [
] No (please explain), [] I don't know

Results

 In the case of negative results, is there a statistical power analysis (or an adequate Bayesian analysis or equivalence testing)? [X] Yes, [] No (please explain), [] I don't know

- Are the results described and interpreted correctly? [X] Yes, [] No (please explain),
 [] I don't know
- Discussion
 - Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument? [X] Yes, [] No (please explain), [] I don't know
 - Are the conclusions adequately supported by the results (without overstating the implications of the findings)? [X] Yes, [] No (please explain), [] I don't know

The MS discusses the term "behavioral flexibility" and how to measure it. The authors discuss that currently since it is very difficult to measure, proxy equivalents are used without having previously verified that these approximations are correct. Common proxies used for behavioral flexibility are sociability, use of urban habitats, and eating new human foods. The main idea is that based on a correct estimation of behavioral flexibility, this concept can be used in studies that will include multiple species.

The MS is well written. The introduction presents the background with the defects that the authors want to correct. With this objective, the authors have presented a project that studies the great-tailed grackle (*Quiscalus mexicanus;* hereafter grackle). This species is a good model since it is social and has expanded its distribution by taking advantage of modifies habitats and by eating new foods generated by human activity. The project contemplates capturing wild individuals and later in captivity, measuring and training some individuals to promote their behavioral flexibility. Subsequently, following individuals released into the wild, the authors corroborate whether the commonly assumed predictions are met when using a proxy instead of behavioral flexibility "per se".

Originally, the project involved testing individuals from three populations: one within the species' historical range and the remaining two from urban and recently expanding areas. Finally, and due to justified difficulties, the authors only worked with individuals from the two recently expanded areas. The authors captured individuals in the wild and tested them in the laboratory with two devices that measured their ability to cope with new technics for access to food under conditions of deprivation. Half of these individuals were then randomly assigned to training experiments in which they were forced to reverse learning those new capabilities. Furthermore, the time required to proceed with reversal learning was used as a measure of its flexibility. Finally, when released into the wild, all individuals (tested only and tested plus trained) were followed to measure many characteristics commonly assumed to be indicators. Overall, they corroborated the predicted association between flexibility and breadth of foraging techniques. However, contrary to expectations, the least flexible individuals used the highest proportion of human foods within their ranges, suggesting that they specialize in those types of foods. Finally, the authors found that there was no association between flexibility and social or habitat use behaviors.

Overall, the hypothesis, predictions, and the MS is really interesting. The results are correctly presented, analyzed, and discussed in detail. In addition, when the results did not hold some

predictions, the authors added new analyses. I have read all the models presented and think that they are correct. However, I did not test the models by running them.

To the best of my knowledge, I consider MS to be very valuable and would be of great interest to scientists working in many branches of ethology and comparative biology or behavioral evolution. My comments are mostly of form or related to the presentation of data and results. However, based on some results, I have a personal comment that I think the authors can discuss related to the use of human foods and the quality of the individuals who depend on them. There is a growing literature on the effect of human diet on animal health, longevity, lifetime reproductive success, and population viability. For example, in the house sparrow (*Passer domesticus*) the decline of its populations has been attributed to its expansion into agricultural and urban habitats. Possible causes of such decline included poisoning from the use of agrochemicals, exposure to air pollutants, low reproductive success in modified habitats, and dependence on human foods that have very low nutritional quality (Vincent 2005, Shaw et al. 2008, Seress et al. 2012, Bichet et al. 2013, Morrison et al. 2014, Berigan et al. 2020). Furthermore, house sparrows from a rural population forced to follow urban diet resulted in a decrease in their body condition and nutritional physiology, and the induction of oxidative stress (Bernat-Ponce et al. 2023).

Therefore, I wonder if the high dependence of the least flexible individuals on human food found in this MS may be a consequence of those individuals being the "low quality" ones. As low-quality individuals are not able to obtain all their nutritional requirements from high-quality natural sources because they cannot compete with individuals of higher quality or rank, they are forced to rely on less preferred foods. Expanding that concept to the population level, these low-quality individuals could not survive in the natural environment. Modification of human habitat and food sources only allows those "low quality" individuals to survive and reproduce. In summary, I think it can be argued whether in those species currently expanding their range into human-disturbed or even urban habitats, the individuals that took advantage of those new opportunities could truly be characterized as "winners" and those species and populations as "successful."

References

Berigan, L. A., Greig, E. I., & Bonter, D. N. (2020). Urban house sparrow (Passer domesticus) populations decline in North America. *The Wilson Journal of Ornithology*, *132*(2), 248-258.

Bernat-Ponce, E., Gil-Delgado, J. A., Guardiola, J. V., & López-Iborra, G. M. (2023). Eating in the city: Experimental effect of anthropogenic food resources on the body condition, nutritional status, and oxidative stress of an urban bioindicator passerine. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, *339*(9), 803-815.

Bichet, C., Scheifler, R., Coeurdassier, M., Julliard, R., Sorci, G., & Loiseau, C. (2013). Urbanization, trace metal pollution, and malaria prevalence in the house sparrow. *PloS one*, *8*(1), e53866.

Morrison, C. A., Robinson, R. A., Leech, D. I., Dadam, D., & Toms, M. P. (2014). Using citizen science to investigate the role of productivity in House Sparrow *Passer domesticus* population trends. *Bird Study*, *61*(1), 91-100.

Seress, G., Bókony, V., Pipoly, I., Szép, T., Nagy, K., & Liker, A. (2012). Urbanization, nestling growth and reproductive success in a moderately declining house sparrow population. *Journal of Avian Biology*, *43*(5), 403-414.

Shaw, L. M., Chamberlain, D., & Evans, M. (2008). The House Sparrow *Passer domesticus* in urban areas: reviewing a possible link between post-decline distribution and human socioeconomic status. *Journal of Ornithology*, *149*(3), 293-299.

Vincent, K. E. (2005). Investigating the causes of the decline of the urban House Sparrow *Passer domesticus* population in Britain. PhD Thesis, Montfort University.

More specific comments below

Line 144. P1 alternative 2. I have a doubt. Always the more caloric foods be considered as being more valuable? I am not a specialist on food quality, but I suspect that there may be some king of foods that are more easily digested. Or that they may have specific micronutrients for example. Of course, number of calories may be appropriate proxy for quality. But the authors may discuss their approximation and alternatives in the discussion of the MS.

Is there information on the reproductive success of individuals and populations of grackles from natural or native vs modified and urban areas? Perhaps more modified habitats imply low breeding success and are used by low quality individuals.

Is food from human sources healthy for grackles? See the paragraph where I cite the known examples for the house sparrow.

Line 233. "grackles" no "rackles"

Line 339. You add microhabitat as a random factor because you want to account for differences in microhabitat but you were not interested in measuring the effect? Please clarify.

Line 398. I suspect your Fig 1. previously had a format more similar to a figure. Now I think is more of a Table with schemes in the upper line displaying the apparatus you had used during flexibility trials. Therefore, I think it would be more clear if you cite the Fig 1 as "Table 2" And refers to the apparatus: see, Table 2 upper line "B". This would be more clear if you add a letter (A, B, C, D) to corresponding cells of the upper line in the Table.

Line 413. Idem my previous comment. Replace Fig 1c for: Table 2, upper line "C".

Line 554. I am not familiar with Bayesian analyses. The "89% compatibility interval" is commonly used in Bayesian analysis? Please clarify.