Dear Author,  
  
I regret that I am unable to accept your paper for publication in Animal Behaviour. Both reviewers agree that you have conducted a nice experiment on an interesting question, with potentially interesting results. However both reviewers also make clear that there are major concerns with the experimental design that requiring addressing before the results can be properly considered. In particular, Reviewer 1 points out the major issue of not randomising the order of the experimental conditions, and both reviewers have concerns with the non-independence of personality measures, and the lack of consideration of possible other explanatory effects. Finally, Reviewer 1 suggests that more justification is needed for the use of bullet ants as a predator model, and I strongly agree with them that further tests or further observation is required to justify the use of this particular model in this context.   
  
I believe all reviewers they have provided some very useful suggesting about possible way forwards in this very interesting experiment, and so I suggest you read their comments carefully in best considering how to reformulate your manuscript.  
  
I am sorry not to be able to give you a more positive answer, but I hope our referees' detailed comments (see below) prove useful should you decide to collect extra data or submit your manuscript elsewhere.  
  
Best wishes  
Dr Lucy Aplin  
Editor  
  
  
-----------  
  
Reviewer #1: This study aims to test whether three behavioral traits (exploration, risk avoidance, and arousal) can predict foraging efficiency in long-billed hermit hummingbirds (Phaethornis longirostris) and whether this relationship is modulated by an individual's perceived predation risk. The authors conclude that foraging efficiency is influenced by personality (i.e., exploration and risk avoidance), perceived predation risk, and their interaction. The underlying question is an interesting one, and I strongly feel that studies examining the connection between animal personality and foraging efficiency are worthwhile. This being said, although I respect the authors' efforts, there are several core methodological issues that are a cause of concern for me. I have detailed these, as well as other points, below. I hope that the authors will find my comments helpful.    
  
Major Comments  
Given that bullet ants are not predators of long-billed hermit hummingbirds, I am unclear why the authors chose to manipulate the perceived predation risk of the hummingbirds by creating situations where a dead bullet ant either was (high "predation" risk) or was not (low "predation" risk) present. Bullet ants feed on nectar, and given this, they may be considered a competitor for hummingbirds (which also feed on nectar), rather than a predator. Furthermore, the authors provide no evidence to suggest that bullet ants will aggressively defend a food resource from hummingbirds (i.e., sting their avian competitors), and it may be the case that the hummingbirds prefer flowers without bullet ants in order to avoid competition for resources (note that the authors only provide an anecdotal observation, made by one of the authors prior to the study, that bullet ants deter hummingbirds from feeding). In other words, there is **no evidence to suggest that bullet ants actually present a  
risk to hummingbirds**. Given that this study was conducted to assess how perceived predation risk interacts with personality to influence hummingbird foraging efficiency, it would have been far more effective to use a predator model for this work.     
  
Why wasn't the order of **the high and low risk conditions randomized**? The high risk condition (i.e., with the bullet ant) was always presented after the low risk condition (i.e., without the bullet ant) and the authors found that foraging efficiency was lower in the low risk condition. Nevertheless, the hummingbirds would have higher motivation to feed earlier in the day, and the longer latencies seen in the high risk condition could simply be due to a satiation effect, since the hummingbirds had already fed. If the authors had varied whether the high or low risk condition was presented first, these effects could have been teased apart, but as things stand, it is not possible to do so. I am, therefore, not confident in the conclusion that the presence of the bullet ants (i.e., the author's measure for predation risk - but see my comment above) resulted in decreased hummingbird foraging efficiency**.** I understand that nothing can be done about this now, but if the authors could  
present evidence to show that, under normal (i.e., non-experimental conditions), there are no temporal effects on foraging throughout the day, this would certainly help increase my confidence in this aspect of the study.  
  
I find the **nonindependence of the recorded personality measures and the measure of foraging efficiency** to be a bit problematic. The authors used personality measures recorded during the foraging efficiency trials to predict foraging efficiency, but these are all part of a single trial. Most notably, the authors use the "total foraging duration" to calculate foraging efficiency (the response variable) and to calculate exploration (a predictor variable). If the authors wanted to assess how animal personality predicts foraging efficiency, I feel that it would have been much more appropriate to conduct an isolated personality assay(s) for individuals (perhaps testing individuals singly in a lab setting) and observing how the resulting measures predict foraging efficiency in the described setup.  
  
Were there situations where more than one individual was feeding at the same time? If so, it may be wise to consider this in the analysis given that **social interactions** can alter an individual's behavior during personality assays, for example, via social conformity (Webster & Ward 2011; King et al. 2015). This is one reason why individuals are often tested individually for personality.  
References:  
1)      King, A. J., Williams, L. J., & Mettke-Hofmann, C. (2015). The effects of social conformity on Gouldian finch personality. Animal Behaviour, 99, 25-31.  
2)      Webster, M. M., & Ward, A. J. (2011). Personality and social context. Biological reviews, 86(4), 759-773.  
As a side note, I will mention that personality can also influence social behavior (e.g., Aplin et al 2014; Firth et al. 2018).  
References:   
1)      Aplin, L. M., Farine, D. R., Mann, R. P., & Sheldon, B. C. (2014). Individual-level personality influences social foraging and collective behaviour in wild birds. Proceedings of the Royal Society B: Biological Sciences, 281(1789), 20141016.  
2)      Firth, J. A., Cole, E. F., Ioannou, C. C., Quinn, J. L., Aplin, L. M., Culina, A., ... & Sheldon, B. C. (2018). Personality shapes pair bonding in a wild bird social system. Nature ecology & evolution, 2(11), 1696-1699.  
  
The authors should have a **paragraph defining animal personality and discuss its relationship to foraging** in the Introduction. A number of studies have demonstrated links between animal personality and foraging behavior and should be cited (e.g., Aplin et al 2014; Patrick & Weimerskirch 2014; Harris et al. 2020; Jeffries et al. 2021)  
References:  
1)      Aplin, L. M., Farine, D. R., Mann, R. P., & Sheldon, B. C. (2014). Individual-level personality influences social foraging and collective behaviour in wild birds. Proceedings of the Royal Society B: Biological Sciences, 281(1789), 20141016.  
2)      Harris, S. M., Descamps, S., Sneddon, L. U., Bertrand, P., Chastel, O., & Patrick, S. C. (2020). Personality predicts foraging site fidelity and trip repeatability in a marine predator. Journal of Animal Ecology, 89(1), 68-79.  
3)      Jeffries, P. M., Patrick, S. C., & Potts, J. R. (2021). Be different to be better: the effect of personality on optimal foraging with incomplete knowledge. Theoretical Ecology, 1-13.  
4)      Patrick, S. C., & Weimerskirch, H. (2014). Personality, foraging and fitness consequences in a long lived seabird. PloS one, 9(2), e87269.  
  
I am a bit concerned about **pseudo-repeatability (also known as pseudo-personality)** in the current study, given the fact that four sessions were performed over ca. two weeks (as a side note, please give the exact timeline here - i.e., provide the number of days instead of saying "ca. two weeks") which equates to an assay every ca. 3.5 days. To avoid pseudo-repeatability, the intervals between assays should be relatively long compared to the lifespan of the species in question. An inter-assay interval of 3.5 days may be a bit short for this species. Nevertheless, I think that a quick acknowledgement of this in this Discussion would suffice. For further reading on this topic, please see Dingemanse and Wright 2020  
Reference  
Dingemanse, N. J., & Wright, J. (2020). Criteria for acceptable studies of animal personality and behavioural syndromes.  
  
More information is needed about the models that were run. For example, I don't understand how the authors got a sample size of N=192. It appears that the authors only studied 11 individuals (according to Table 2), and if each individual was assayed 4 times in the high and low risk conditions, this should equate to 2\*4\*11= 88 = N, not 192. Is this because more than one foraging event was counted for a single individual per day for each risk scenario? If so, I might recommend using only the first visit for each individual under each risk scenario for each day instead. Furthermore, it is unclear what other variables (besides perceived risk and the 3 behavioral measures) were included in the model. For example, site really should be incorporated into the model as a fixed effect (note, that site only has 2 levels and must be included as a fixed effect and not a random effect) given that the authors ran this study at two different sites. **Individual ID should be included as a  
random effect**. **Date (or days since the experiment started) should probably also be included as a fixed/random effect to account for any effects of seasonality.** It is unclear whether any of these were included in the models.  
  
At multiple points throughout the discussion (e.g., Lines 233, 242, 274) the authors suggest that their behavioral measures (i.e., exploration, risk avoidance) \*may\* be related to personality, but by definition, given that these behavioral traits are significantly repeatable, they \*are" measures of personality.  
  
In terms of readability, I would strongly recommend the use of an **English proofreading service**. While grammar/spelling issues are unrelated to the quality of the underlying science (and should therefore have no bearing on the acceptance of this manuscript, in my opinion) there are overarching issues that will need to be addressed before this manuscript can be published.   
  
Minor Comments  
Lines 52-54: Just say that there is a continuum of personality types and remove the bit about "polar-opposite phenotypes".  
  
Lines 64- 67: "Under viarable conditions, a diversity of behavioral strategies can be mantained when not a single one performes in an optimal way in all contexts. If so, performance may vary as a function of the interactions between social or environmental selective forces" A quick note that the social environment is part of the environment, so I might change this to read "social or physical environmental conditions". References are also needed here (e.g., Bergeron et al. 2013; Le Cœur et al. 2015; Roth et al. 2021; Mouchet et al. 2021; but see Roth et al. 2019).  
References  
1)      Bergeron, P., Montiglio, P. O., Réale, D., Humphries, M. M., Gimenez, O., & Garant, D. (2013). Disruptive viability selection on adult exploratory behaviour in eastern chipmunks. Journal of Evolutionary Biology, 26(4), 766-774.  
2)      Le Cœur, C., Thibault, M., Pisanu, B., Thibault, S., Chapuis, J. L., & Baudry, E. (2015). Temporally fluctuating selection on a personality trait in a wild rodent population. Behavioral Ecology, 26(5), 1285-1291.  
3)      Mouchet, A., Cole, E. F., Matthysen, E., Nicolaus, M., Quinn, J. L., Roth, A. M., ... & Dingemanse, N. J. (2021). Heterogeneous selection on exploration behavior within and among West European populations of a passerine bird. Proceedings of the National Academy of Sciences, 118(28).  
4)      Roth, A. M., Firth, J. A., Patrick, S. C., Cole, E. F., & Sheldon, B. C. (2019). Partner's age, not social environment, predicts extrapair paternity in wild great tits (Parus major). Behavioral Ecology, 30(6), 1782-1793.  
5)      Roth, A. M., Dingemanse, N. J., Nakagawa, S., McDonald, G. C., Løvlie, H., Robledo‐Ruiz, D. A., & Pizzari, T. (2021). Sexual selection and personality: Individual and group‐level effects on mating behaviour in red junglefowl. Journal of Animal Ecology, 90(5), 1288-1306.  
  
Lines 82-83: Unclear.  
  
Line 94: Please include **a paragraph detailing the ecology** of long-billed hermit hummingbirds.  
  
Lines 98-99: I am confused about what N was. Here the authors state that "Of the 21 birds marked at the study lek, 12 individuals regularly visited commercial hummingbird feeders" yet Table 2 suggests that 11 individuals were studied. Please clarify.  
  
Lines 106-107: Did each phase last from 3-4 hours, or was this the total time? Please clarify.   
  
Line 123: How far away from the feeders did an individual have to be to consider it within the "feeder area". i.e., at what distance did data recording begin?   
  
Lines 151-152. Please describe how models were fitted using the rptR package.  
  
Line 162: Please report the variance inflation factors.   
  
Lines 179-180: Please give Repeatability Estimates and p-values/confidence intervals in text.   
  
Line 191: Given that this was conducted at two different sites, site should be included as a fixed effect in the models.   
  
Lines 229-234: I am confused given that lines 147-148 suggest that exploration was only measured during the low risk treatment. Also, it is not explained in the Statistical Methods section that the authors examined how predation risk predicted exploration.  
  
Line 237: In the Methods the authors do not state that they measured food intake so this comes out of nowhere.  
  
Line 262: What are trapliners and territorials? I'm not sure this paragraph is necessary. It appears to come out of nowhere.  
  
Line 272: Individual identity was not discussed as being included in the models in the Methods. This was also not mentioned in the Results.  
  
Line 287: The Acknowledgements section has multiple ellipses. Please complete this section.  
  
  
------------------  
  
Reviewer #2: Previously, studies on the relation between foraging behaviour and risk allocation did not include the effects that individual differences might have in this interaction. Recent studies, however, have found that indeed individual differences do play a significant role in explaining when will an individual forage despite a possible risk.    
The authors argue that hummingbirds in general, and the long-billed hermit hummingbird in particular, make a useful model to study the relation that individual differences have in foraging behaviour and risk sensitivity. The authors compared the foraging efficiency of marked wild hummingbirds feeding from an array of feeders with and without a simulated predation risk. They found that indeed the hummingbirds' foraging efficiency was lower when there was a predation risk present and that different measurements of foraging behaviour differ between the two treatments but were repeatable within individual birds. While the number of feeders used and the amount of movement made by the hummingbirds were significantly affected by whether there was simulated risk present at the feeder, the latency to approach the feeders did not change with the treatment. The authors discuss how hummingbirds changed different foraging behaviours in relation to the risk treatment and how different  
bird "personalities" might have affected the results found.   
I do recommend the study to be published after a major revision since the subject and findings should be of interest to the readers of Animal Behaviour but there are some problems with the clarity of the manuscript in its present form. The experiment was well planned and the results are clear but some additional analyses would make the study more compelling. I do have some general suggestions and detailed comments that I hope will improve the manuscript.  
  
General comments:  
The introduction could be clearer in explaining **the rationale, specifically how the risk allocation hypothesis relates to behavioural plasticity or indeed personality**. A concrete example of the relation between different personalities and foraging behaviour would be useful for the reader. Predictions could be more clearly layout for each of the behaviours measured. For instance, **authors could clearly state that the degree of exploration (feeders visited) is expected to decrease in the presence of the simulated risk but that the birds that explored more during the control treatment were also expected to explore more when a risk was introduced compared to hummingbirds that explored less during the control treatment**. The introduction is not clear in whether the three behaviours measured were supposed to vary between individuals or correlate to each other.

**Methods are clear but lacking some detail**. For instance, how many birds were tested in total? The methods say 12 birds visited the feeders regularly but then Table 2 has an N=11. In any case, it should be mentioned how many birds were tested in both treatments (with and without the risk) and how many birds were tested for each of the four sessions. Did all the birds complete all four sessions? How many times did individual birds visited the array in each of the treatments and was there any competition between birds visiting the array? Was there an effect of the number of times individual birds visited the feeders and the foraging efficiency under either treatment? For the statistical analyses is not clear that the repeatability test included "individual bird" as the grouping variable or what error distribution was used for that analysis. Figure 1 is useful and well designed. "Feeders design" needs an apostrophe.

Both the results and discussion section could be improved if authors were to include more information about the distribution of the different foraging behaviours in response to the risk treatment. While there is individual repeatability in most of the behaviours measured, it is not clear whether there were distinctive strategies (behavioural syndromes) between the hummingbirds. This would be interesting and complementary to the results already presented. It is also not clear, at least to me, that the behaviours measured correlated to each other. Did the birds that approached the array of feeders faster also visited more feeders and were more efficient?

Finally, individual differences might be the result of differences in motivation driven by physiological state, sex or other factors that need not be explained by hummingbirds having different personalities. I think this could be acknowledged. Given that in this species of hummingbirds sex can be readily determined by the bill morphology, sex too could be included as an explanatory variable. The same could be done with the weight of the birds as a proxy for motivation, but only if the data was collected. While there is no doubt that there are behavioural differences between the hummingbirds feeding in two different risk contexts, is not clear that those differences are consistent between the different behaviours recorded and that differences could not be explained by other factors.   
  
Detailed comments:  
Lines 4: Should read "Increasing number of studies demonstrate high variation in foraging behaviour"  
  
Lines 6-7: The sentence is hard to read, please rewrite.   
Lines 7-10: From this sentence, it could be interpreted that the authors tested fitness consequences of foraging behaviour, which is not what the authors did. Please rephrase.  
Line 38:  There is no need for acronyms in this text.   
Line 39: change to: "..prey allocates time for foraging inversely proportionate to…"  
Lines 48-50: Should read: "Existing evidence demonstrates not only the variation per se but consistent inter-individual differences in the average level of a behaviour displayed across a range of contexts"  
Line 57: remove one "in"  
Line 63: Should read: "The fitness payoffs of a behaviour are expected to drive" The word payoffs does not need hyphenation.  
Line 69:  There is no need for acronyms, use the full name, "the hummingbirds" or "the birds" instead.  
**Line 75: No need for author's name initial  
Line 78: No need for author's name initial**  
Line 86: Are foraging spots the same as feeders? If so, then change to feeders or keep consistent throughout the text.   
Line 108: No need for author's name initial  
Line 115-116: Please provide more detail. Did individual hummingbirds vary in how often they visited the feeders?  
Lines: 116-117 Change or rephrase to "Two exemplary video clips of the two phases of the experiment, are provided in Supplementary Materials (video 1, 2)".  
Lines 147-148: Were these behaviours also not measured during the high-risk treatment?   
Line 179:  "Repeatability" is in bold letters, it should not be.   
Lines 211-212: Please rephrase.  
Lines 219-220: Was there an effect of experience? Was the risk avoidance lowered after the first time a hummingbird visited the feeder array with the ant?  
Lines 244-246: Please rephrase, the sentence is confusing.   
Line 248: Change to "the role of predators in the  evolution of personality"  
Line 272: The plural of analysis is analyses  
Lines 308-309: The citation might be incorrect.  
Lines 317: Italicize the Latin name  
Lines 338-340: The reference is repeated  
Lines 353-354:  Please italicize the Latin names  
Please make sure all references have the same formatting throughout  
Figure 3 **could be improved by adding a box plot on top of the violin plots.**  
---------------  
  
Reviewer #3: This is an ethical review. I have no major ethical concern. However**, I would encourage the authors to expand their ethical note.**  
Animal capture and handling for marking purpose: the authors should provide more details about how the birds were captured and comment on potential adverse effects of the capture and handling procedures.   
  
Foam tag weight: The weight of the foam tag should be documented, in absolute term but also relative to the weight of the individuals. The authors noted that only 12 individuals over the 21 marked birds visited the feeders. Is there any indication that the tags might have increased mortality by preventing some individuals to forage normally and/or increasing predation risk?  
  
Regarding the experimental design and the manipulated level of predation risk, I do not have any ethical concern: the authors used a dead ant (found dead in the forest) and individuals had the choice to not expose themselves to the perceived threat after the first visit (other sources of foods were available).  
  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
In compliance with data protection regulations, you may request that we remove your personal registration details at any time.  (Use the following URL: <https://www.editorialmanager.com/anbeh/login.asp?a=r)>. Please contact the publication office if you have any questions.