

## Response Letter

Thanks for evaluating our manuscript. The reviews were really helpful to us and led us to considerably rework the MS. We are particularly satisfied with our extended results regarding the “Death all or nothing”, which is now unified with the other Death scenario. This is a major improvement in our Results that helps see the unifying aspect of the rMVT, and a novel section was added to present the rMVT as one entity, not a collection for particular scenarios. Also, to address the reviewers comments we have refocused the Results section, in a way that should help readers understand what practical consequences are (beyond the theoretical construct). Below are our point-to-point replies (in blue).

Dear Authors,

The article that you requested be evaluated has been read and commented on by two reviewers and myself. All the readers found the subject interesting and the potential contributions interesting. There are several areas where readers asked for more detail and clarification, and questioned whether all of the technical details had been properly considered. One reader also felt that the overlap with existing publications was high, and that the authors needed to explain more clearly how this current work fits into the broader literature.

I was very interested to read this article and found the introduction very helpful for getting an idea of where the article was going and of summarizing existing work in the field. I have a few specific queries of my own related to the derivations and results:

1) the use of the phrase "transformation" and of "newtonian" time seem unnecessarily confusing. The functions presented are expected gains or expected amounts of time spent. There is no actual transformation of time. I would simply call these functions expected gain and expected amount of time spent in patch.

We agree that this vocabulary was somewhat non-standard, especially for biologists. We have much streamlined the terminology. In particular, we've abandoned the term “Newtonian timescale”, which was probably the most facultative and physics-oriented term. When needed, we now use “regular time” instead. We've had to retain “transformation” at several places, since it describes well the fact that the x-axis (and the y-axis) must be transformed (as a log-transform would) in order to be able to use the rMVT method. However, we no longer use “transformed timescale” to refer to those axes, and use instead “expected time” and “expected gains” when appropriate, as you suggested.

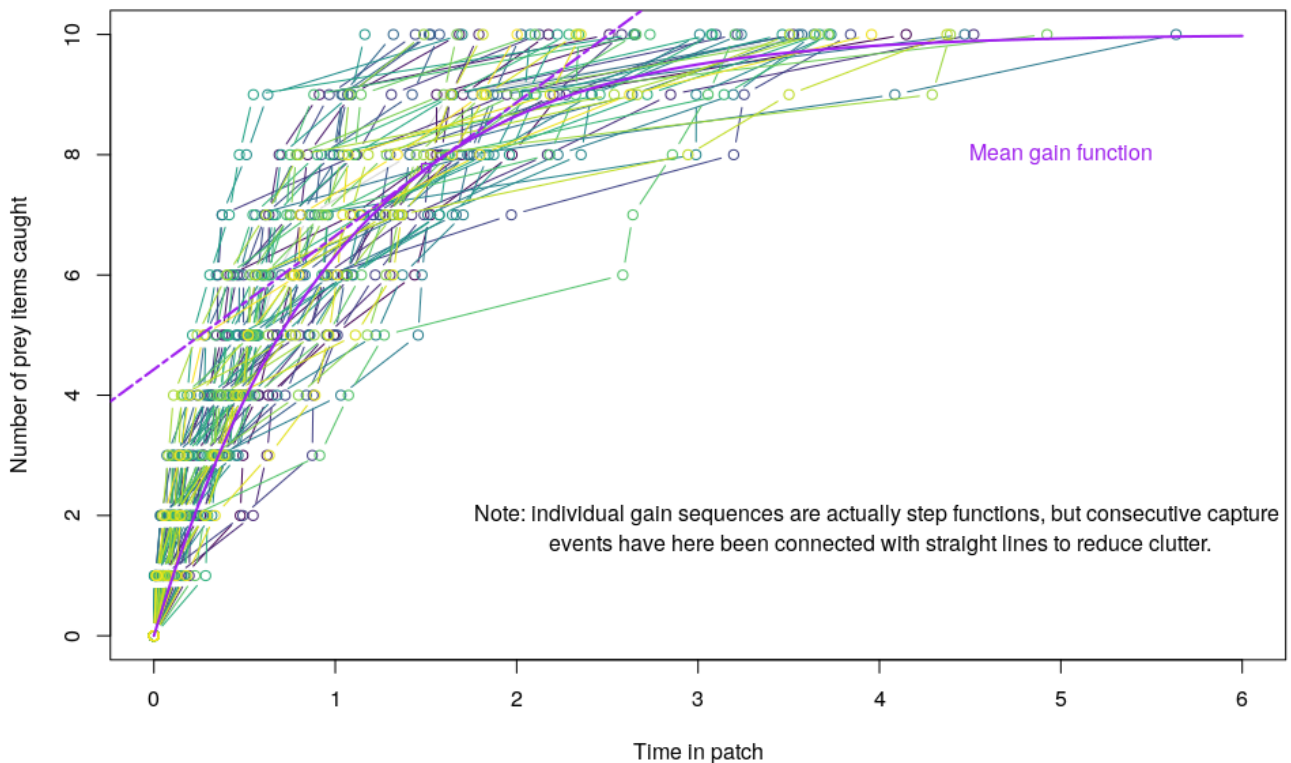
The relevant “timescale” (what used to be called the transformed timescale) is not always simply the “expected foraging time”; it is not so in the Death scenarios, where micromorts must be used. So we now use “optimization domain” as the general term, to encompass “expected time” and “micromorts”. Finally, we have added a new Section “The rMVT in a nutshell” to unify the different particular scenarios into one framework, and this section also helps clarify the terminology. The beginning of the Results section, dealing with the particular scenarios, was made shorter, to make room for this novel unifying section. We think this new section really helps the reader see the big picture of rMVT and get the terminology, while also addressing the remark from Reviewer 3 that the different scenarios were individualized and not integrated in a common framework.

2) The article focuses on calculating expected gain as a function of the strategy, and the strategy is the amount of time spent in the patch. I don't see derivations in the article for the optimality of the strategies. In particular, equation 2 defines a function  $F_{hat}(t)$  that represents the mean energy gain for departing after  $t$  time units. While this equation makes perfect sense to use when calculating the average gain per time one gets when leaving a patch, it actually does not represent the current gains within a patch. If a focal animal has not been disturbed and reaches  $t^*$ , it will have a different level of gain than the derivative of the expected value of gain. I do not see an explanation in the article for how using the derivative of expected gain gives the optimal solution. It seems to me that a strategy that leaves when the current gain drops to the expected average gain is optimal, but this does not result in a strategy of leaving at a constant time.

You make an important point, one that involves the interpretation of the MVT, and so you'll excuse us for providing a pretty lengthy response here. First of all, let us clarify that the possibility you raise for an individual to decide to leave a patch at a certain time or rate is only conceivable in one of the 6 risky scenarios: the "Distraction" scenario. In other cases, an interruption forces the individual to leave (or die); it has no choice. The individual can aim at leaving after some time  $t^*$ , but it may well be forced to leave before, and has no control on when this happens. This effect changes the marginal rate at which an individual will decide to leave if un-interrupted (by changing the overall rate of gain), but does not mean it will actually leave patches when this target rate is attained (as it may be forced to leave earlier). Under the "Distraction" scenario, we focus on relatively frequent interruptions, with relatively mild and short-term consequences, such as the distant presence of a predator causing a forager to raise its head at each suspicious move. Less diffuse cases in which an interruption has severe and long-lasting consequences, to the point that leaving after one is the obvious thing to do, are covered in the second scenario ("Escape" scenarios).

This proviso made, you raise the question of whether optimality means an individual should decide to leave when its current (instantaneous) rate of gain in the patch falls below the average rate of gain in the environment. Whilst this is a commonly stated prediction of the MVT, it must be remembered that it's a prediction, not a postulate. The only postulate is that "*the individual leaves at a time such that the long term average rate of gain is maximized*", and this is the postulate we use. It turns out, in some conditions, that this is achieved if leaving a patch when the marginal rate equals the average rate. However, this prediction holds strictly only in the case where the acquisition of gains in a particular patch obeys a smooth, concave, gain function, with no stochasticity. Although this assumption is commonly made for illustration, it need not be in practice.

Let us illustrate this in the most classical example of MVT applications: a predator is hunting for prey with some attack rate  $a$ , and a patch is characterized by the number of available prey items in it ( $n_0$ ). In this case, the gain function used is the exponential  $n_0(1 - \exp(-a t))$ , from which one derives the optimal residence time (or giving-up time) and/or optimal GUD (Charnov 1976). Yet, in any particular visit of a patch, the accumulation of gains follows a stochastic sequence of prey captures, producing a step function. It looks like this:



This plot help make several things clear:

-1- The MVT does not predict a departure time for each particular stochastic realization: it only predicts one optimal strategy for the average gain function (the exponential curve in gray), i.e. for **the average gains expected over all particular realizations for this class of patches**. The fact that the optimal residence time is when the instantaneous (marginal) rate of gain equals the average rate of gain applies for the average (smooth) gain function only.

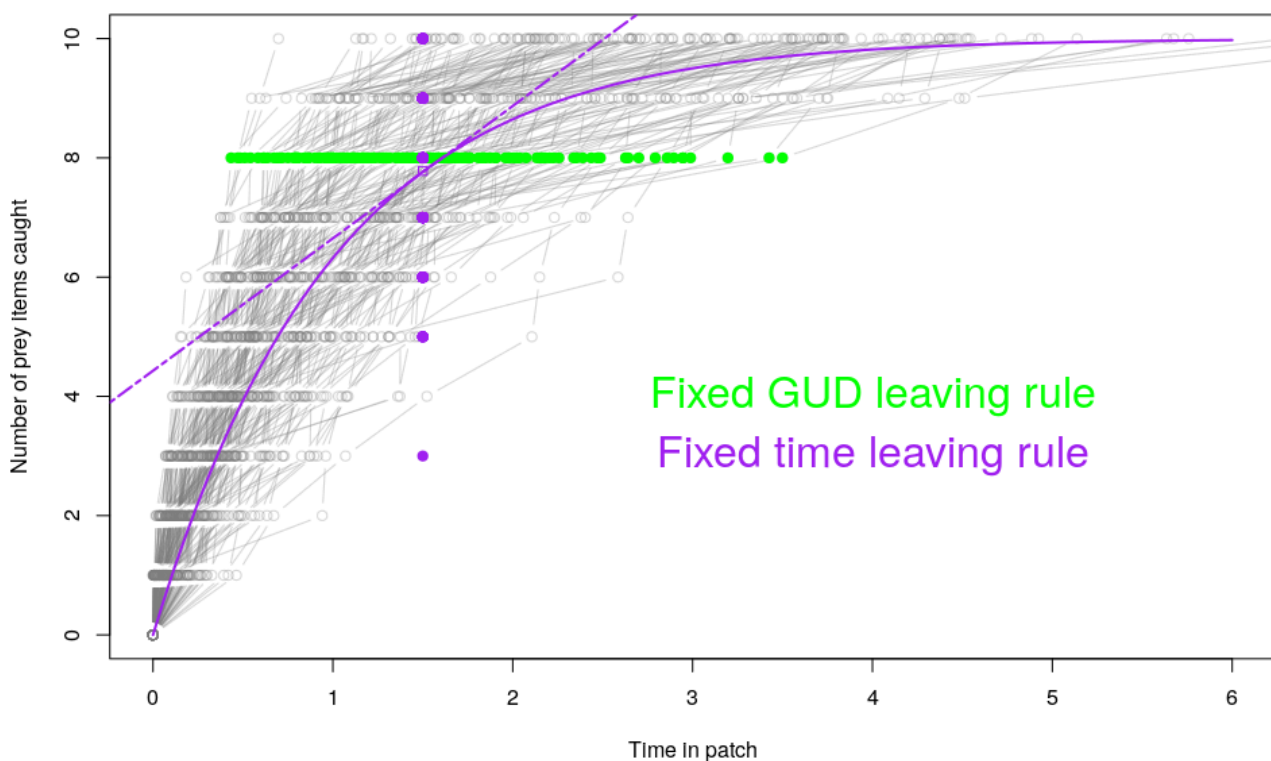
-2- An individual does not leave when the instantaneous rate of gain **for a particular visit** falls at the average rate. First, one cannot properly define an instantaneous rate of gain for a discontinuous function. Second, if we tried to smooth every step function and apply this reasoning, it may yield disastrous results: the individual could persist too long in some “lucky visits” (e.g. on the top left), or leave too early in initially “unlucky” visits (bottom right). This would not correspond to the MVT, and would be sub-optimal. One strategy that is close to attempting this is the so-called “giving-up time (GUT) rule”, which is indeed suboptimal (not shown here; but see e.g. McNair 1982).

Points -1- and -2- show that the MVT predicts an optimal residence time for a given **class** of patches, which is characterized by a smooth gain function, encapsulating the expected gain over the ensemble of possible particular visits in that class of patches. (This is, btw, one reason why the MVT assumes that there should be a large enough number of patch visits. And why experimental tests average over several patch visit replicates).

Our approach is the same as the standard MVT case illustrated above, except that the stochasticity in particular gain functions does not result from stochasticity in prey capture, but from randomness in the occurrence of interruption events. In the same way as the MVT, we use as a basis for prediction the average gain function for a particular class of patches (the function you are referring to). As for the usual MVT, the equality between the average rate of gain and the marginal rate of gain does hold (see all the figures with the tangent constructs), though of course it holds only on the appropriate time scales and gain functions (those defining

the rMVT). Just like for the usual MVT, during a particular patch visit, an individual would not always leave when its instantaneous rate of gain falls below the average (an instantaneous rate of gain may be hard to define, or would be very fluctuating, owing to stochasticity).

One final remark: the specific behavioral mechanism through which an individual implements the MVT (or other) optimal strategies (referred to as patch-leaving rules) is unspecified. Many possible alternatives or “rules of thumb” have been discussed (counting elapsed time, updating internal motivation...), and the rule “leaving when the rate of gain falls below the average rate of gain” is only one of them, albeit not a particularly realistic one (Stephens & Krebs 1986; McNair 1982). That particular mechanism should not be regarded as defining the MVT: the MVT simply predicts optimal residence times (and corresponding GUDs). As an illustration, we show two alternative ways (both achieving very similar performance in this example) an individual may implement the MVT optimal strategy: (1) by leaving as soon as the prey density falls below the GUD  $n^*$  (GUD rule; in green), and (2) by leaving after a constant resident time  $t^*$  (time rule; in purple).



In the first case (GUD rule), the individual would leave at different times at every particular visit; in the second (time rule), it would leave at different prey densities in different particular visits. In both cases, it would achieve the optimal strategy predicted by the MVT, yet in neither case would it be leaving when the instantaneous rate of gain equals the average rate. The latter property holds only in average, for the ensemble gain function (purple curve).

We hope this addresses your interrogations. We have added a paragraph in the Discussion, as well as more explanations in the beginning of the Methods, to explain these aspects. We have also added an opening Section in the S.I. (“General assumptions”) to present (in condensed form) the elements given in our answer above. Finally, the results now present both optimal residence times and optimal GUDs, and compare the two in relation to patch leaving-rules, which directly addresses the matter.

3) After equation 3, it seems that the described scenario involves an animal losing time if it is disturbed. But it can just start foraging in another patch, it doesn't lose the time. This is only revealed much later in the paper.

After eq (3), the individual escapes the patch when disturbed indeed. But it does not immediately start foraging in another patch, since it first needs to find that other patch. Thus it does lose some time ( $T$  units of time, i.e. the average travel time between patches). We now write this explicitly in the beginning of that "Escape" section (line 175), to avoid any possible confusion.

I believe there is a critical piece missing from this article and that you need to show how the strategy of leaving when the expected gains per time drop to the global expected gains per time. Other than that, the main recommendations involve altering how you talk about transformations and non-newtonian time and some better integration of existing literature.

Thank you for your careful reading. We hope the changes we made and our explanations address these concerns.

## REVIEWER 1

I very much enjoyed this manuscript. It attempts to unify two previously separate subfields, a goal I consider valuable and important. The paper opens with excellent motivation -- even though I am not a specialist in either MVT or Brown's GUD frameworks the authors convinced me that unifying the two is work worth doing.

Much of the mathematics was well explained, the results appear sound, and I appreciated the intuition they provided behind many of them. Because I am not a specialist I will focus my review on large picture issues and narrative flow. I did not check many of the technical mathematical derivations (my stochastic theory is embarrassingly bad).

[Thank you for your positive comments on our manuscript and the many suggestions!](#)

### Major Issues

I see only a few "major" issues. I suspect none are fatal, but only parts in need of clarification. However, some of these parts are for core ideas of the paper and so I am calling them "major."

### MVT Graphical Analysis and concavity of the gain function

The paper often mentions that one can apply the same graphical analysis as in MVT. My understanding is that these graphical methods only work or give unique solutions if certain technical conditions hold (ie. the gain function being concave or some similar shape that you can find the tangent line). It is not obvious to me that the transformed gain functions meet these requirements. I would like to see more discussion on this and proofs that the graphical analysis can always be applied.

[This is a good point indeed. We now provide mathematical justifications that the gain function  \$G\$  matches the conditions of applicability in a new Section in the S.I. \(S. I. Section 4\) and indicate this in the Results \(lines 281-282\).](#)

This graphical argument appears similar to that of the Levin's fitness function/set method, which is also used in the Smith-Fretwell model. It is an adjacent field, but still might be worth citing. [It is indeed related, even though the graphical argument of the MVT is more specific \(it involves the notion of a travel time, not just the existence of some long term rate of gain\). In other words, the method allows one to find the slope of the line, it does not require it as an input \(in contrast to more general optimization models\). For this reason, we opted not to cite the other optimization models you are mentioning.](#)

The paper left me unsure of how important is it that the gain function for each patch is concave? It seems to me essential to the argument as a few derivations rely on Jensen's inequality. I would like to see expanded discussion around this, whether non-concave change the results, and whether such functions exist in nature.

[We have clarified the conditions of applicability of our approach in the new opening S.I. Section \(S.I. section 1\). Gain functions need not be strictly concave in every patch, they can have in principle the full generality allowable under the MVT \(as explained in some detail in Calcagno et al. 2015 and 2017\). For the graphical method to hold in particular, restrictions are the same as for the usual MVT.](#)

The Jensen's inequality argument (which is, more precisely, a Delta-method argument; see S. I.) works irrespectively of the sign of curvature, even though the conclusion that it will cause the expected gains to decrease holds only in the concave portions of the function (which would be the largest portions or, at least, the ones involved around the optimal strategies). We have reformulated this part to clarify.

#### Additive partitioning

(Line 296) Is additive partitioning provably impossible to achieve? Or were the authors unable to find a way? I'm reading it as the former but if so I would like to see the proof. I imagine the proof is fairly trivial, but as this is an important claim I think its worth writing down in the supplementary materials.

Yes it is. This is now made more explicit: we have added a paragraph explaining the argument and stating what forms of risks can or cannot be formulated in "Brownian" form, and why (page 10). We also refer more extensively to the S. I. sections where we show the derivations and why we can/cannot reach a Brownian form. We also spend more time explaining the common use of Brown's equation in the main document (lines 288-292).

The whole distinction between Charnovian and Brownian formulations and what exactly is meant by the partitioning are also worth discussing further. I would like the authors to expand on the significance that only one case can be written in Brownian form. What does this mean for this larger program of unifying Charnovian and Brownian views? What does this mean biologically?

This is a good point. Our motivation was not to undermine or promote one or the other theory, but to see that risk can be incorporated into the MVT framework. But we understand one naturally wonders what are the implications at the end of the day. The article is overall now much more oriented towards showing predictions and their biological meaning, and much less to theoretical constructs and considerations. We highlight predictions that could not be obtained from Brown's GUD theories, several of which are for scenarios that cannot be put in Brownian's form. We think this shows how it can be useful to have a more versatile tool.

#### Mathematical Terminology and Notation

I found the terminology and notation to be fine overall, but I ran into certain problem areas which I think could be improved for clarity.

I felt the term "transformation" along with the tilde notation and  $F$  are overloaded with different meanings. In equation 2 we have the "transformation of the gain" and in equation 4 we have the "transformed gain." Are these different names meaningful? I thought not on first read-through as they both use the same notation. But then it becomes confusing on whether 4 is building on 2 or is a completely separate case.

My current understanding is that  $F$  in equation 2 is the "real life" gain function when we add in risk, escape, and other details. That is, it is a better approximation of reality than the original gain function  $F$ . I feel this use of the term "transformed" conflicts with "transformed time" and the transformed gain of equation 4, both which feel more of a technical mathematical manipulation. We now explain in much greater details how, starting from an initial gain function  $F$  (the one we would use in the classical MVT with no risk; eq. (1)), we may obtain a different function that describes the expected gains, and eventually an effective gain function  $G$  (the one to be used in the rMVT, on the effective optimization domain). We have changed both our notations ( $\tilde{}$  was indeed overloaded with different meanings) and our terminology (we now distinguish "expected effective foraging time", "expected residence time", "effective optimization domain"...), without using transformed as a catch-all term. We hope this is clearer now.

In equation 5 we have this dot notation defining the function. I have not seen this notation before and I'm confused about how the function is defined. My current reading is that it is typoed and means  $G(\cdot) \equiv F(\cdot)$ . Furthermore, it seems that  $F$  is called the gain function,  $F$  is the transformed gain, and  $G$  is back to being called the gain function. This is confusing.

The dot notation for the argument of a function is commonly used to reduce clutter, though it is not indispensable (the dot means "whatever the argument of the function is"). We have abandoned it in the first equation for  $G$ , and we now explicitly write that function  $G$  is a function of  $\tilde{t}$ . We just retained the dot function the second time  $G$  is introduced (in Death scenarios), since the equation is then inline and reducing clutter is more important. We now call function  $G$  the effective gain function (i.e. the one to be used eventually in a rMVT), and explain in greater detail how it is obtained from several transformations of the initial gain function  $F$ .

Sometimes the patch subscripts are dropped. I am confused on what this indicates.  $\beta$  is a particular offender. I could not find where it was defined yet it is of importance later one, especially when defining optimal boldness.

Subscripts (indicating different classes of patches) are dropped when there is only one class of patches (homogenous habitat). This is the case we focus on in the examples, for simplicity. This is now stated explicitly (line 90, then lines 333-334, and also in the new Table S1).

The paper uses risk-MVT and rMVT interchangeably. Sometimes both versions are used in the same paragraph (eg Table 1 caption) making it appear they refer to different things. I would choose one and toss the other. (My vote is to keep risk-MVT and it is clearer and still fairly concise.)

We have harmonized our uses within sentences and paragraphs. We now recall for each section that risk-MVT is abbreviated as rMVT. We retained the use of rMVT, since we use it many times and a few spare letters end up making a difference. Also, "MVT" was an acronym, and so rMVT keeps the same spirit of being an acronym.

Overall I would recommend the authors revisit some of this confusing terminology, and provide a table of notation. With the six different cases there are enough moving parts that a table will go a long ways towards readability.

We have followed your recommendation and now provide a Table of notations and symbols (Table S1).

### Optimal Boldness

I found the distinction between boldness and risk confusing, and whether these are per-patch or global definitions. It seems that boldness to me is  $\beta t$  while risk is  $\beta$  (which I remark again is undefined!). But in that case isn't optimal boldness is just the optimal foraging time rescaled? Once you know the optimal time  $t$  you gain no new information by considering optimal boldness. It feels more like a definition, one which I do not understand the significance of.

We have now clarified what  $\beta$  means, which will probably help here.  $b^*$  is intended as a global definition. In that section we focus on homogeneous habitats, so the issue does not really apply.

But we have added the corresponding expression for  $b^*$  when patches have different characteristics, which should also help understand that it is a global concept.

Regarding " isn't optimal boldness is just the optimal foraging time rescaled?": in some sense, yes it is! But it should be remembered that foraging time  $t^*$  is a nonlinear function of  $\beta$ . We might as well write  $b^* = \beta t^*(\beta)$ . This shows why when we vary the "riskiness" of patches ( $\beta$ ), it is not trivial to predict how  $b^*$  will vary, as it all depends on the quantitative response of  $t^*$  to  $\beta$ . We hope this clarifies the fact that  $b^*$  is not a trivial concept, and brings some additional information not contained in  $t^*$  alone. We now explain this in the main text, and also explains more why this metric allows us to better understand the behavioral adjustments occurring when risk increases..



To me optimal boldness has a connotation of deciding how risky of a patch you are willing to visit. I.e. something the forager controls separate from time spent. By the importance given to optimal boldness in the title and abstract I expected it would be a deeper and more discussed aspect. Instead it seems quickly gone over and insignificant. I would like to either see boldness have a deeper treatment, or for the title and abstract to change to deemphasize this. Personally, I think some title that makes it clear the authors are unifying Brown's and Charnov's frameworks would be better, as that seems to be the deep significant result and the soul of the paper.

We agree that optimal boldness arrived a bit as an afterthought in the paper, which reflected the chronology of our research. However, we do believe it is an important concept that one can address with the rMVT, and a concept closely linked to current theoretical and empirical questions in ecology and behavior. For this reason, we opted for the first of your options: giving optimal boldness a more thorough treatment, with more explanations, and more references.

Regarding your view on optimal boldness, we think this is just a matter of being discrete/qualitative (do I or not visit this patch?) versus continuous/quantitative (how much risk do I accept to take in this patch?). But the underlying concept is the same. One strength of the MVT over, say, diet selection models, is that it allows for gradual responses (how much do I exploit this patch type?) versus binary responses (do I exploit this patch or not?). And binary responses are a subset of gradual responses: some patches might be totally rejected under the MVT ( $t^* = 0$ ), which effectively represents the fact of “not being willing to visit” a patch. This can happen for too poor patches (or too risky patches in the rMVT). In fact, the MVT can be shown to generalize simple yes-or-no diet selection models. This is why we think that our  $b^*$  metric is a quantitative metric of boldness, just like  $t^*$  is a quantitative metric of patch use (not just yes-or-no, but “no” definitely being a possibility).

### Small things

Here are some minor points. Some of these are important but easy fixes. Others are small nitpicks. Many are likely due to my unfamiliarity with the subfield, so I defer to the authors whether these are worth changing.

(160) I'm likely being dense, but I do not see how this is a linear rescaling. Did the authors mean it is approximately linear for high values of  $\beta$ ? I would appreciate clarification.

It is a linear rescaling of time because, in eq. (2), in the first term  $t$  is replaced with  $(b/(g+b))*t$ , which is time multiplied by a constant. In other words, the average effective foraging time is proportional to the residence time  $t$ . Since we now put less emphasis on timescale transformations and have streamlined these sections, the sentence involving “linear rescaling of time” has been removed.

I'm confused about the distinction between transformed time and effective time. Would be nice to define what a micromort is within the text. I do not think this is familiar to many readers, but seems easy to define.

A definition of micromorts is provided on page 8, lines 252-256, with references. We now explain in greater detail the distinction between effective time, expected time, and the “optimization domain” (which used to be called “transformed time”). The new Table S1 should also help.

"It is quite universally true in biology that organisms do not perceive time uniformly, but rather in a relative way, depending on internal state and external factors." Is it? It sounds believable but it is still a strong claim and I would like more citations.

Indeed. In addition to the example of thermal biology, we have added a reference to differences in the context-dependent perception of elapsed time in primates. We also slightly deemphasized the statement.

#### Equation 1

The MVT is specifically saying that the solutions of equation 1 maximize individual fitness, correct? I found this a little unclear. Perhaps putting line right before the equation can help.

It is correct. We have modified as suggested.

(75) It would be nice to mention what the  $n$  in  $E$  stands for to help mnemonically. Later on I forgot and kept thinking it was an index term like  $j = 1, \dots, n$ .

We've done as suggested.

#### Figure 2

The quotation marks in (b) both face the same direction. (Actually, this is true for all figures.)

Fixed.

I would appreciate slightly more detail for the caption of (b) since it could serve as a nice reference for what each of the six elementary combinations mean. Alternatively, we could have a table reference for the combinations. I really liked how the authors considered these six combinations as they seemed very natural, but I'd like the reference as I found some of the names difficult to remember.

We agree that these names may be hard to remember, and they are slightly arbitrary (different authors have been using different names). We now present them as a Table, as you suggest.

#### Paragraph 131

Is there any reason to call these states (i) and (ii) rather than the foraging and interrupted states? The latter is easier to remember and read and I see no advantage on the former.

Agreed. We have changed as suggested.

I am curious to why the return to foraging rate  $\gamma$  is not based on patch type? It seems a reasonable simplification but I would like just a line or two of justification.

It could be. In this article we focus on the simple case where all patches have the same parameter (we thought of this as a property of the individual rather than the patch). But you are right this parameter could vary as well. We have modified accordingly.

Does  $\delta$  scale with how long the interruption lasts? Or is it a one off payment for being interrupted. I'm reading the former in the equations but the text is slightly ambiguous.

As parameterized,  $\delta$  is seen as a one-off payment. We have clarified.

#### Figure 3 + 4

(Figure 3) The formatting on the plots seems to have exploded.

This should be fixed.

(Figure 4) What do the blue lines mean? Is this meant to show the one-to-one mapping?

Yes. This is now much more explicit in the revised Figure 3.

(Both figures) What do the grey lines mean?

The grey lines are the tangent construct that allows to solve the (r)MVT problem. We have added this in the caption.

#### Figure 5

Is the "all or nothing" blue line covering the orange? Or does the effective timescale not exist here?

This Figure has been entirely replaced.

(Table 1) Could the caption clarify what is meant by rMVT domain?

It now does (as well as the main text).

Figure 5 typo "I each case..." I found three or four other typos two I forgot the location of. All minor, but maybe worth a copy edit pass.

The Figure has been reworked entirely.

I found the introduction well written, clear, and concise. The conclusion I felt was less so. It's not bad, but the quality of the introduction spoiled me and I suggest another revision pass on the conclusion. I listed many issues in the name of constructive critique, but I want to emphasize again that I greatly enjoyed this manuscript and the ideas in it. It is an excellent paper with a significant contribution. I hope others like it as much as I did.

Thank you. We've much reworked the Conclusion part; we think it is better now.

## REVIEWER 2

One reason that GUDs and MVT receive separate treatment is that they are pretty different - one (GUDs) is more of a qualitative description of a phenomenon, equation notwithstanding, but that offers a method for uncovering certain values experimentally, as Brown has done with squirrels and other organisms; the other (MVT) is a theoretically-motivated equation that permits much more rigorous mathematical analysis and extension.

This is interesting, but we are not sure this perception of things is universal. Brown's framework is based on equations, and it can make quantitative statements (e.g. Brown et al. 1992). The fact that it is derived from equations is what allows us to determine when its predictions are applicable (or not). Conversely, the MVT is (most) often used qualitatively: "*the residence time is longer when the travel time is longer*" is a qualitative prediction, and arguably the prediction most often tested. Another prediction is "*Stay longer when the average habitat is poorer*", which is also a qualitative test commonly done. Last, the MVT is often presented graphically, without showing the underlying equation (see King & Marshall 2022 *Curr. Biol.* for a recent example). Our take is that both approaches are equation based, both can yield quantitative or theoretical predictions, and both can also yield qualitative easy-to-test predictions.

Putting them into a common framework is commendable and reasonable, but does not lack precedent, which the authors do allude to (though there are many additional intermediate developments over past decades that are not referenced here). In addition, a recent paper which is cited throughout this one (Arehart et al, 2023) explicitly unites MVT and GUD approaches, at least in the "death" scenarios of the present work. While there is clearly room for additional development in this space, I don't think it is clear what the authors' new approach offers that is particularly valuable compared to previous contributions in the area.

As we say in the abstract and the article, we seek to retain the simplicity and graphical appeal of the MVT. Approaches like the one by Arehart et al. (2023), and more generally stochastic dynamic programming approaches, are optimal foraging approaches tackling a similar question, but they are quite different (and much more complicated/computationally intense), than the other two. We say this explicitly in the Conclusion, and the revised version is refocused on predictions and practical consequences, which should help see what is the novelty.

Perhaps this is just a matter of presentation. The main part of the paper reads a bit too much like description of the derivation of the math, rather than providing clear connections/examples relevant to applications. The rescaling of time in particular is something that makes sense, but could make sense much more quickly if it is explained differently. An ecologist without expertise in this specific realm of theory would, I feel, likely get lost/lose interest unless more examples are identified and presented. For instance - what is an example of a distraction vs costly vigilant scenario, amongst real organisms' foraging strategies? What about the escape examples? You highlight the example of a bee or reproducing organism for the "all or nothing" scenario, but could present that in more depth, and compare to other scenarios for which your framework might be useful.

Your point about the rescaling of time and its explanations is somewhat reminiscent of the Editor's comment on the vocabulary: we have changed our terminology and we no-longer use "transformed time-scales" or "Newtonian time", in favor of more simple terms, as recommended by the Editor. We agree that your general impression: the paper was quite theory-oriented. We have much shortened the mathematical parts, and refocused the Results Section on actual biological predictions, with the addition of five new references to the empirical literature..

I do think there is value in making some sort of distinction between (a) fitness dynamics due to behavior in response to perceived predation risk, and (b) the fitness dynamics of actually dying because a predator kills you. However, I found the breakdown into six categories a bit contrived.

If the purpose of the paper is to provide a unifying link between GUDs and MVT, then a more straightforward and intuitive interpretation would be helpful. Is it really illuminating to make categorical distinction between "disturbance" and "escape," or would it be sufficient to present the disturbance scenarios as a subset of escapes? Within disturbance, isn't distraction just a cost-free subset of costly vigilant scenarios?

We guess this is largely a matter of presentation indeed. It seems to us natural to distinguish "disturbance" and "escape" scenarios, as they correspond to quite different biological situations, and also they have quite different theoretical consequences, as we demonstrate in this article. Within disturbance scenarios, the two extreme cases can indeed be seen as part of a continuum (this is explained on line 165), but it is simpler to contrast two extreme cases. The same thing could be said within Escape scenarios (a forager might lose  $x\%$  of its gains when escaping a patch), and possibly within Death scenarios (a forager might convert  $x\%$  of its earlier gains into fitness when dying). But making all scenarios special parametric cases of a giant continuum would make derivations more complicated and perhaps more confusing. Brown's GUD-approach, for instance, took one specific scenario only ("Death all-or-nothing"). That said, we have added a new "The rMVT in a nutshell" Section where we explain how to combine all the different scenarios into one general rMVT, and provide explicit guidelines on how to use it.

In addition, it strikes me that these three categories of scenario must all be relevant for any real foraging organism. Are there cases where a forager would have to be costly-vigilant, but would not potentially encounter one of the death scenarios? How does the framework integrate those? Presenting a worked example incorporating all of these elements would be invaluable, and demonstrate the new exhibility a orded by this approach.

This is a good suggestion. For clarity, we have taken the usual approach of treating different limit cases where one clearly identifiable example dominates the others ("pure cases"). In reality, many different processes might be combined, making everything more complicated. This is now discussed explicitly in a new Section (page 9). We also provide an example where two scenarios are mixed (Fig. S2) and discuss what practical implications this could have (Conclusion).

When drawing the connection between rMVT equations and Brown's GUDs (Pg 9), the authors only draw a clear connection between 'death' scenarios and GUDs. This is also demonstrated in the Arehart et al paper, so I feel it would be necessary to clarify the distinction between that paper's derivation and the present one. It also feels a little like the paper has overpromised by offering a synthesis between GUDs and MVT, but only providing it in these specific cases.

Our result is precisely that there can be no connection in several cases (the "non-Brownian" ones; Table 1). In those cases, the equation underlying Brown's GUD result just does not apply. The fact that in some specific conditions introducing risk in the MVT does yield Brown's result is already known indeed (this point was made by Brown itself), but the new results are all the other possible cases. This is now made more explicit (see also our answer to the question from Reviewer 1).

Overall, I think that this work could form a useful contribution to the MVT/GUD/foraging theory literature, but unfortunately I believe it needs to be rewritten to better highlight the innovations and place them in a more intuitive context.

We have made profound changes in the Results, Conclusion and Figures, and added some additional references, to that effect.

#### MINOR NOTES:

The paper does contain some typos, and would benefit from a careful editing pass.

We have done so.

The title of the paper appears as "TAKING FEAR BACK..." on the title page, but as "WHEN CHARNOV MEETS BROWN" on subsequent pages - which is it?

“When Charnov meets Brown” is the so-called running title, different from the full title (not all journals have one).

Figure 2: I would consider redrawing this as a table, rather than as a graph, as I feel like the implied axis scales could be misleading.

We have done as suggested.

Figure 3: the lines are drawn well outside of figure margins, creating a confusing plot.

Fixed.

Figures 4-5: There is a lot of emphasis here on comparing the true and effective timescales, which I think could serve to confuse ecological readers, and could be more fruitfully replaced with more figures exploring the dynamics of examples from the 6 scenario types

These figures have been replaced with new Figures more focused on ecological consequences.