phencc
Note that reviewer comments are in *italics*, while our responses are in regular text, and all in-text citations generally cross-reference to the main text.

Editor's comments:

reviewers were quite critical of a number of aspects of the article. In the end, I think the biggest issue is one of communication. The authors need to focus their arguments much more clearly and deliberately.

We appreciate the editor's comments about clarity of message, and agree it does much to explain many of the reviewers' concerns. To address this, we have overhauled the manuscript, especially sections 1.1-1.3 (now sections 1.1, 'Defining & measuring tracking' and 'Tracking in single-species environment') to be more precise and shorter. We have overhauled our figure defining tracking and been more careful in our definition of fundamental versus environmental tracking. Additionally, we have tried to more clearly separate evolutionary and ecological theory, which we believe understandably led to some confusion. We believe the revised submission is much improved and explain our changes in more detail in our point-by-point response to reviewers below.

We appreciate
Referee 1 comments:

The authors present a manuscript that attempts to summarize our current knowledge about ecological tracking, i.e. the ability of an organism to track the phenological niche. This is particularly interesting in the context of climate change and earlier onset of seasons in the northern hemisphere. The topic of phenological shifts is interesting, and I found the manuscript overall very well written.

We thank the reviewer for the positive comments on the manuscript's topic and writing style.

I have a few general concerns about the manuscript which I detail below, and some specific ones, which I will address later in a chronological order.
1) I am not familiar with the topic of ecological tracking, but I am very familiar with the literature regarding phenological shifts in response to climate change. In my opinion, ecological tracking appears to me as a rebranding of a phenomenon about much has been written. I am aware that the authors will disagree with this view, but their manuscript did not convince me that ecological tracking is fundamentally different from the widely observed phenological shifts. Maybe it is a subset of those, but it is nothing new. Nevertheless, the effect of phenological changes on ecological communities is an interesting one.

We agree our manuscript's topic is most easily and readily applied to phenological shifts (as Reviewer 4 also noted), but we avoided this term given that phenology is generally defined as 'the recurring timing of life history events' (defined on line phendefine1) and a number of events we review (and to which this manuscript applies) fall outside this definition. We do understand that definition of phenology may be evolving in the literature and have tried to up front about the reasoning for our terminology now, when we define tracking we now state (line phendefineS-line phendefineE):

> Both these definitions are readily applied to phenology—the timing of recurring life history events—though they can also apply to non-recurring life history events (e.g., seed germination), or events not normally defined as part of life history.

2) After carefully reading the manuscript, I did not understand what this manuscript actually is about and what the authors want to achieve with it. a) The authors claim it is a review, but many studies- and many reviews about them (e.g. by C. Parmesan or A. Menzel) - have described phenological shifts in response to climate change. Only very few of those are mentioned, and in the description of their narrow search criteria they end up with only a handful of studies, because it appears that the reviews and the studies therein were actively omitted.

This is a good point, as we were too broad in our previous draft of our aims (e.g., 'we review current knowledge on tracking both in empirical data...'), and made it seem we were aiming to review the full literature on phenological shifts. This was not our aim, and we are more specific now (see line aimS, "Here, we review the concept of tracking used in the empirical climate change impacts literature and through its related ecological theory."

Regarding the references—we did cite (Menzel *et al.*, 2006) and now also cite (Parmesan, 2006). As our aim is not a full review of all studies of phenological shifts we did attempted to balance older and newer references, as we wanted to balance a mix of foundational studies with new work that gives up-to-date estimates based on to-date climate change (i.e., studies from the mid 2000s generally use data when global average warming was much lower).

b) It is also not clear to me why they reviewed these papers and not the theoretical literature or the physiological literature. Both types of studies were discussed in detail in the manuscript but not reviewed at least I would doubt that the lack of studies identified by the authors regarding theory or the physiology of the cues can be based on a handful of studies. There must be myriads of studies in animals and plants addressing the physiological basis of cueing for phenological events, e.g. flowering time in plants or breeding time or migration time in birds. I was particularly surprised that they also excluded theoretical studies in their search, while at the same time relying heavily on theoretical papers throughout the remaining manuscript to describe several aspects of ecological tracking and its consequences for populations and communities. If this was a review, why exclude theory?

We believe the reviewer is here (and below) referring to the part of the paper regarding a targeted systematic literature review for studies examining tracking and other traits together. This review is only mentioned in the Box 'Trait trade-offs with tracking,' and is not meant to be the focus of our paper. For this review we did not exclude physiological studies, though we did exclude modeling and theory studies because they did not have data (only 12 of 231 papers). We have tried to clarify this in our text within the Box and in the supplement (e.g., we have renamed this section 'Literature review of studies examining tracking & traits,' and we now

open this section with "To examine current evidence of what traits may trade-off with tracking".

We completely agree with the reviewer that theory and physiology are quite relevant to our topic and do review a number of relevant studies throughout the main text.

c) It was unclear to me whether they were searching for studies that explicitly talk about ecological tracking (which are, I believe few), or any study that has ever observed a shift in phenology due to warming. The latter is not achieved, but it is also maybe not needed given the many reviews we already have. The former is probably not needed, too, because ecological tracking is, in my opinion, largely a rebranding of (adaptive) phenological shifts.

Again, we believe the reviewer is here referring to the part of the paper regarding a targeted systematic literature review for studies examining tracking and other traits together, which is only mentioned in the main text in the Box 'Trait trade-offs with tracking.' We have worked in the supplement to clarify that we are specifically looking for studies that examine tracking and traits at once; our search terms do not require the term tracking (or track*) but do require reference to a trait. Thus many studies that only examine phenological shifts would be excluded, as finding those studies was not the aim of this systematic literature review.

d) If I accept it is not a review, then it is possibly an opinion paper or a perspective. I understood that the authors mention a whole suite of understudied aspects of ecological tracking and that they want to fuel a whole suite of new studies. However, for a perspective, the rationale for addressing some of the understudied aspects of ecological tracking is not always clear. For example, for studying mismatches between phenologies of coexisting species, it is not crucial to know the exact cue. Also, while the need for non-stationary models appears logical, I could not find anywhere clear predictions about why and how coexistence mechanisms would be changing differently in non-stationary systems compared to stationary (but fluctuating) ones. This is regrettable because I assumed that the interaction between tracking and coexistence mechanisms was a main focus of this manuscript at least this would be an interesting topic.

We appreciate the reviewer's concern and it is in line with Reviewer 4's concerns as well. To address this we worked to focus more on the interaction between tracking and coexistence mechanisms. To do this we have merged two former sections and significantly streamlined the sections before 'Tracking in multi-species environments.' We have not completely removed these section as we belive (as did previous reviewers in their comments) that some background is needed before the section on coexistence mechanisms. Additionally, we given an example of a model with a fluctuating environment where stationary and non-stationary outcomes are not the same in the Box 'Adding tracking and non-stationarity to a common coexistence model.' Finally, we have updated Figure 2 to clarify why we believe the cues matter.

In fact, I would not expect large differences between a classical storage effect model and a model where the environment changes gradually and directionally over time, especially as storage effect models also look at environments with different statistical properties. Specifically, if say, we have a storage effect model (or a model addressing priority effects) where the environment does not fluctuate strongly, species would probably not be selected for being able to track, simply because tracking is not needed when the environment is stable. However, if we model (as in a classical storage effect scenario, or in a priority effect model) the environment as highly variable and unpredictable in time (and space), then species inhabiting such an environment must be able to

track, because they cannot know what the ideal timing would be in any given year, unless there is a good cue (in which case the environment would not be unpredictable). Thus, I would expect a similar change from non-tracking to tracking when comparing stable with fluctuating (stationary) environments as when comparing a stationary with a non-stationary one. In other words, species inhabiting highly variable environments should be tracking, which may equip them with an advantage also in a gradually changing world. This idea has been voiced before in models (e.g. Bonebrake, T. C. & Mastrandrea, M. D. 2010. Proc. Natl. Acad. Sci. USA 107: 1258112586) but also in experimental studies conducted in fluctuating habitats, where no effect of experimentally induced climate change was found.
So maybe the lack of a prediction about why we should look at non-stationary models and how their outcome would be different from what we know may be explained: the outcome would not be much different. It is also possible that the authors had attempted to exactly derive such a prediction in their model in the previous version of this manuscript, but I understood that they did in fact not produce any surprising results.

We give an example of a model with a fluctuating environment where stationary and non-stationary outcomes are not the same in the Box 'Adding tracking and non-stationarity to a common coexistence model.'

Help from Megan needed!

3) I was also not sure what exactly the topic of this manuscript is. From the previous reviews and the authors replies I understood I that this manuscript aimed at coupling ecological tracking theory with coexistence theory, which would be an exciting topic. However, only approx. 10% of the manuscript is devoted to this topic. The remaining 90% are spread across several different and partly unrelated aspects of ecological tracking. These are, to name a few, the lack of physiological evidence for cueing, definitions of ecological tracking and measuring it, description of bet hedging as opposed to tracking, a brief note about the equivalence of phenotypic plasticity and ecological tracking, trade-offs between tracking ability and competitive ability (why this trade-off and no other one?), and some more. Interestingly, none of these various topics in actually reviewed in detail, which brings me back to my initial question of whether or not this is a review.

We appreciate the reviewer's concern and have worked to streamline the manuscript so that more of the text is devoted to'Tracking in multi-species environments.' Sections on physiological evidence for cueing, definitions of ecological tracking and measuring it, and review of plasticity versus bet-hedging are now shorter, but we have not removed them because we believe they are critical background for examining ecological tracking and coexistence, and highlight areas where we need advances if we hope to better understand tracking and coexistence. Previous reviewer comments also stressed these connection and we think they are important, but we could have done better to present more briefly and as background, which we now do.

In my opinion, the authors do themselves a disservice by evoking expectations about linking ecological tracking with coexistence theory, when in the end they spread sometimes thinly across several aspects of ecological tracking. The manuscript could thus realty profit from being concise in the selection of aspects discussed and then discuss these aspects exhaustively.

As mentioned above, we have worked to streamline the sections outside of those in tracking in multispecies environments.

4) It is not clear to me why out of all possible biotic interactions, competition is dealt with so prominently. I understand that competition is the other side of the coexistence coin, but since coexistence theory is not the core of the manuscript, other biotic interactions should have been discussed, too. There could be positive interactions that are decoupled by climate change and (as mentioned by the authors) decoupling of interactions among trophic levels. The subsequent focus on trade-offs between tracking and competitive ability appears to me equally arbitrary. If we accept that plasticity comes at a cost, it can trade-off with any trait. For example, I would think that stress resistance (which in plants is assumed to trade-off with competitive effect ability) would trade-off with tracking, ability, too. Also, there could be trade-offs between phenological plasticity (i.e. tracking) and plasticity in other traits that enable fitness homeostasis even if no ecological tracking occurs. This relationship is not addressed. However, it could be fundamental if organisms are highly plastic in other traits, in which case they may not even need to track.

We focus on the trade-off between tracking and competitive traits as it is predicted by theory and the most supported by empirical evidence.

5) Ecological tracking is regarded exclusively as a plastic response. However, the (very few) solid studies on evolutionary change in response to climate change indicate that phenological traits could be among the first under real selection. I was asking myself why plasticity should be the main mechanism by which species can track, and whether we need this assumption for defining ecological tracking, or whether the definition could also embrace rapid evolutionary change.

We understand the reviewers concern that adaptive tracking (*sensu* Simons, 2011) theoretically could equally explain tracking and we understand the concern that there not many enough rigorous studies on evolutionary change in response to climate change. However, most studies we are aware that have estimated plastic versus evolutionary change in phenology find it is mostly due to plasticity and many phenological traits are highly plastic (if the environment is defined in calendar time) thus we have retained our focus on plasticity but now have worked to be clear about this, line itsnotevo, "Given our focus on responses to climate change, we consider environmental tracking here as a mainly plastic response (Bonamour *et al.*, 2019), though over longer timescales or in certain systems it should be shaped by selection (Franks & Hoffmann, 2012)."

Specific comments (chronological order, line numbers are references given):
Line 1-12: reference to the many studies and reviews about escape in time is missing (e.g. Parmesan, Menzel, and many more). This leaves the impression that we know nothing about ecological tracking, which is, in my opinion, not true.

We now cite Menzel *et al.* (2006); Parmesan (2006) on line r1ass.

37ff: Do we need to show that tracking is related to fitness? Isnt that self-evident and if not, why?

We have removed this line, but have worked to address this in section 'Defining & measuring tracking.'

74ff: Is it true that we know nothing about environmental cues? I did not take the time to dive deep into the literature but I would think that studies on birds and plants are plentiful. Maybe the mechanistic studies (i.e. experimental) are rarer than correlations (but they do exist, e.g.

reciprocal transplant studies and not only Arabidopsis), but even evidence for correlations of e.g. flowering time with e.g. growing degree day units is abundant.

In streamlining the manuscript we have removed this paragraph.

84ff: The advancement in phenology by certain numbers of days has been demonstrated by C. Parmesan or A. Menzel (and others) much earlier than what is cited here. I am puzzled why their work is not cited.

We now cite Menzel *et al.* (2006); Parmesan (2006) on line r1ass1.

93ff: Why is it so crucial to know the exact physiological mechanism of tracking and why the cue? For example, if we are mostly interested in the same trophic level and competitive interactions, we may, as a first approximation, assume that the organisms use a similar set of cues. Also, if it is true that we know nothing about the relationship between physiology and the cue, this seems a rather bleak perspective and may lead to the conclusion that we will never understand ecological tracking. So why is this important?

We have updated Figure 2 to clarify why we believe understanding the cues is important, but have otherwise worked to shorten this section to address this reviewer and reviewer 4's concerns.

192-194. Some variable environments do provide cues, e.g. in the Sonoran desert annual system (see Pake, C. E. and Venable, D. L. 1996. Ecology 77: 1427 1435), the amount of the first rainfall in a year seems to partly predict the rainfall of the season. Predictive germination has also been addressed from a theoretical perspective by Cohen (1967) and subsequent authors.

We agree and cite papers by Venable and which build on Cohen's work throughout the manuscript (e.g., line r1ass2, line r1ass4).

195ff: One important aspect of the cueing seems to me the reliability of the cue. Unfortunately, the authors do not mention this and only focus on benefits and costs. To me, this seems a key aspect which is tightly related to the costs (i.e. low reliability, high potential costs). The reliability is not touched upon in the cost-benefit discussion.

We agree cue reliability is important (though we follow , in considering it a constraint), we have updated Figure 2 to clarify this.

208ff: The discussion about bet hedging is too much black and white (i.e. between not germinating and germinating). There is also plasticity in germination rates and some of it is driven by cues (see literature about predictive germination). I would actually assume that in the classical bet-hedging system (desert annuals), tracking ability would be selected for very strongly because in a fluctuating environment, plants need to respond very plastically to the ever-changing conditions. So the idea that there is either tracking or bet-hedging is not plausible for me.

Agreed, we have re-written the section on bet-hedging (line bhS-line bhE), which now includes, "Environmental variation often includes both predictable and less predictable aspects. In such cases theory predicts organisms may evolve tracking that is a mixed strategy between bet-hedging and plasticity."

217-229: This paragraph does not appear to contain much information, so it could be left out.

We have shortened this into one sentence that we include regarding contraints and plasticity, line r1consS-line r1consE.

243ff: I am missing an in-depth discussion about plasticity, i.e. the ability to maintain fitness (fitness homeostasis) even when the environment fluctuates strongly. Plasticity is expected to evolve under unpredictably varying conditions, and tracking is only one aspect of that plasticity. There should be trade-offs among the different types of plasticity.

We have revamped the section on plasticity (line plasS-line plasE) and worked to shorten it. Given this reviewer and reviewer 4's request to focus the paper more we have kept this section short.

1.4: This paragraph is entirely devoted to tracking-competitive ability relationships. It seems logical that tracking ability should also trade-off with tolerance to stress (e.g. low temperatures if e.g. bud burst is early) which in turn may trade-off with competitive ability.

We believe the reviewer means that tracking could co-vary with stress tolerance, which we agree with and now mention on line r1stress.

336ff: Isnt the storage effect the same as tracking only that it is about inter-annual variation and not variability in intra-annual timing? So what would then be the fundamental difference between stationary and non-stationary models when, e.g. we start with a storage effect model in a randomly fluctuating environment where species must already be able to track? I feel it would be crucial to provide clear predictions about what non-stationary models may predict in contrast to classical models. Without these, the call for more and different models is not very well justified. Here, the main justification is that it has not been done, but not this is why stationary models are entirely misleading. Unfortunately, the Box remains vague about this.

Megan, can you write something here?

1.5 I found this section somewhat if not completely- redundant with the sections before and was not sure why it is needed. Much of the discussion here remains somewhat vague. The conclusions are that we need more interdisciplinarity, more understanding and measuring of tracking, more looking at trade-offs with selected traits, and more models that are different from the current ones. Overall, this is not the strongest section of the manuscript. It could be merged with the previous sections and made much more concise.

We appreciate the reviewer's concerns. We have shortened the previous sections so that this section is less redundant.

Box
578-581: Could the finding of early species tracking more simply be due to the fact that response to environmental variables (e.g. higher temperatures) follow a logistic curve where the late species attain high fitness because they are always in their climatic comfort zone? Whereas the early species experience, during their life or evolutionary history a much larger range of

temperatures, some of which are clearly decreasing fitness?

This is an interesting hypothesis and possible, but we are not aware of any formal studies of this.

600ff: Many models and data have been published about within-season timing of (germination) events. They could make a valuable contribution to this section (e.g. Simons, A. M. 2009. Proc. R. Soc. B 276: 1987 1992. Simons, A. M. 2011. Proc. R. Soc. B 278: 1601 1609).

We agree and now cite line simonsref1 in our section on evolutionary theory. This box is focused on one particular model (an ecological model with no evolution) and for clarity we mention only the relevant model in the Box. Throughout the manuscript we have also worked to clarify where we are speaking mainly about evolutionary versus ecological models.

607ff: I believe that a similar storyline could be created with stress tolerance instead of competitive ability.

Agreed, we focus here on competitive ability as that is what the literature has found evidence for.

Referee 2 comments:

The resubmitted paper How environmental tracking shapes species and communities in stationary and non-stationary systems by Wolkovich and Donahue deals with environmental tracking, specifically how environmental tracking can be measured and analyzed, how it may influence species co-existence and species responses to climate change. I think the topic of the paper is novel and highly relevant, and overall the authors did a very good job in reviewing the literature on the topic. I specifically like the part about how tracking may trade-off with other traits (e.g. those related with competition) and thereby shape the co- existence among species in ecological communities.

We thank the reviewer for their comments and have worked to retain the better parts of the manuscript while improving the rest of it based on feedback from this and the other reviewers.

I only have one point to criticize: although the authors highlight that researchers are increasingly recognizing the need to consider multiple climate variables (L 14) this review is mainly focused on environmental tracking in response to temperature changes. I am aware that there is much more known about phenological responses to temperature change compared to precipitation change, which is also supported by the result of the literature search in the Supplement. However, as this review deals with climate change and not only climate warming and we know that climate change is complex and multivariate, I would love to see more examples in the text about environmental tracking and precipitation change. Are there any studies about how temperature and precipitation change may interactively affect environmental tracking (e.g. via changes in snow cover)? If not, I think this could be highlighted in the future directions paragraph more explicitly. Just out of curiosity, would it be possible to include such interactive effects of multiple resources in the model?

We appreciate this comment and completely updated Figure 2 to address it, working to show that both temperature and precipitation are likely critical for many organisms. We also now mention megadroughts and pluvials (line r2precip2) and have altered our final sentence (line r2precip3). We mention snowpack in model Box (line r2precip) and do believe it could be addressed in the

model by developing a more complex environment and cue system. On evolutionary timescales this question is addressed somewhat in some models, for Chevin & Lande (2015), which we now cite.

L 502 Not only temperature is rising but we already and will face non-homogeneous but fundamental differences in the precipitation regime around the globe

Good point, we now write "in the altered climates of our future" line r2precip3)

Referee 3 comments:

In a review piece, Wolkovich & Donahue comprehensively present the idea of environmental tracking by species in stationary and non-stationary environments. This review is loaded with information and touches on several fundamental ecological ideas in relation to environmental tracking by species. The effort therefore is commendable with a potential to motivate new research avenues for climate change ecology-particularly the phenology research. Having said that, I also struggled at various places to grasp the core idea authors were intending to communicate. I outline them below.

We appreciate the reviewer's time and comments to improve our manuscript. We agree that our previous draft was perhaps so loaded with information that the most important and salient points were lost, and we have worked to fix that as we outline below.

I definitely agree with phenology as a trait and tracking as a plasticity of this trait (lines 244-246). I also liked how authors relate the idea of subsequent trade-offs in traits owing to costs associated with plasticity. I, however, missed examples of which traits and plasticity of them are going to trade-off the most with tracking, and how these may differ in stationary and non-stationary environment. Can we also say something whether the strength of trade-offs may differ in these two environments?

XXX

Difference in species ability to track environmental changes as something similar to competition-colonization trade-off is further a stimulating idea (lines 273-280). I was, however, left guessing if authors modelled this at all in their theoretical frameworks. My initial impression was that figure 3 gets at this, but I am not really sure if two species scenarios in figure 3 relate one species as a competitor (lower cue) and the other as colonizer (higher cue). Can this be clarified or if possible implemented?

XXX

Line 5 (Abstract): species responses

We have changed this on line r3misc, which we hope is the requested change. We have also added line numbers to the abstract to help with identifying the exact change requested.

Line 12 (Abstract): through the lens of which ecological theory? Later, you mention community ecology theory. Perhaps, use the latter to be consistent.

Done, line r3misc1.

Line 2: Perhaps, use more recent IPCC citation.

We believe this is is the most recent citation from IPCC Working Group II ('Impacts, Adaptation and Vulnerability') that considers various warming levels and a full report on impacts. We now also cite the more recent report focused on 1.5 C of warming; if the reviewer is referring to another report, please let us know.

Lines 10-12: The "indirect effects of climate change" is not very clear. Why could it not be a direct effect of climate change? Please clarify.

Good point, we have changed to fitness consequences (line r3misc2)

Line 21: Can you elaborate which foundational ecological theory is meant here?

This was unnecessarily vague; we have changed to 'community assembly theory' (line r3misc3).

Line 43: Which basic community ecology theory? Please be specific when mentioning a theory as you did in lines 23-26.

Done, we now write "community assembly theory—especially priority effects and modern coexistence theory" (line r3misc4).

Lines 237-240: Would not this be a trophic mismatch case still predictable from the stationary environment? Or does this imply that trophic mismatch will not occur in the non- stationary environment? Please clarify.

Good point, we have tried to clarify this without adding too much text. line r3birdsS to line r3birdsE now reads:

> For example, consider an organism's whose cues evolved based on a correlation between peak prey abundance and daylength: the daylength cue that could be reliable in a stationary environment (generally predicting preak prey abudance based on daylength, with some interannual variation), but would become unreliable if warming advances peak prey abundance. Predicting the outcome of non-stationarity would be possible from the stationary environment in this case given researchers know (1) the full cue system of an organism, (2) how it relates to fundamental tracking, and (3) how both that cue system and the underlying fundamental model shift with non-stationarity.

Lines 254-256: But what about the benefit side of the tracking? And which other traits those be where trade-off with tracking will be higher?

XXX

Line 309: two thes

Fixed (line doublethe).

Lines 386-388: Please use this example as a separate sentence.

Done (line r3misc5).

Lines 408: Please provide more example studies when you suggest "many studies".

Done (line r3misc6).

Line 416: Here one or two examples will help the readers.

We have added one example and can give more if requested (line r3misc6S-line r3misc6E).

Referee 4 (Ally Phillimore) comments:

There are some very interesting ideas in this review on phenological responses to environmental change and I can see it making a stimulating contribution. However, there are a lot of aspects that require attention, including the structure. In general I found the ms rather imprecise in its use of terminology and quite verbose. I hope the comments below are useful in revising the ms. I have not really commented on the coexistence theory aspect as I am not sufficiently familiar with this literature.

We are glad the reviewer thinks this piece could make a stimulating contribution, and agree that there was room for streamlining and conciseness, and greater precision in our language. We have worked to address these issues and explain them in more detail below.

My biggest criticism of the ms is that the term environmental tracking, which is central to the ideas being developed is not clearly defined, despite having a section devoted to its definition. A clear definition is provided for fundamental tracking, but then the text switches to environmental tracking without providing a definition (except in fig 2). This term seems to be applied more loosely to any case of phenological change, but initially without any discussion of what the yardstick is (Visser and Both 2005), meaning that its unclear that tracking is taking place for instance the response could be maladaptive. The yardstick for tracking could (from hardest to quantify to easiest) be the rate at which (i) the optimum is changing (as in Chevins B or the authors fundamental tracking); (ii) a resource is shifting or (iii) the environment is changing (Amano et al. 2014). Related ideas are introduced from line 100, but you might consider introducing them sooner. Overall I found sections 1.1 and 1.2 quite muddled. I think environmental tracking as used in these sections is synonymous with how the existing literature would refer to phenological responses to cues (line 109), and I don't see that introducing new terminology brings something useful to the table unless there is also some discussion of how much the environment is shifting, i.e. we need to know something about what is being tracked. Another concern is that introducing new poorly defined terms will just generate greater confusion in the field.

This was also a concern of Reviewer 1 and something we have struggled with. One thing we struggle with is how broad the definition of phenology needs to be to include the diversity of

events we include in the paper and to which we believe the topic of the paper applies. We have tried to clarify this in several ways. We have changed the title to be more specific without (hopefully) jargony and we now try to address this head-on when we define tracking—we now state (line phendefineS-line phendefineE):

> Both these definitions are readily applied to phenology—the timing of recurring life history events—though they can also apply to non-recurring life history events (e.g., seed germination), or events not normally defined as part of life history.

We now provide a new Figure 2 to clarify our definitions and we have overhauled the text where we define environmental tracking (line definetrackS-line phendefineS):

> Conceptual and theoretical treatments of tracking often relate how well an organism matches the timing of a life history event to the ideal timing for that event, what we refer to as 'fundamental tracking'. In contrast, empirical studies of tracking often focus on estimating a change in the timing of an event per unit change in an environmental variable, something closer to what we refer to as 'environmental tracking'—the change in timing of a major biological event due to a species' cue system given change in the environment.

We have then restructured this section to contrast fundamental tracking and 'environmental tracking,' which agree with yardsticks (i) and (iii) of the reviewer. We further clarify what we mean by environmental tracking (line moretrackS-line moretrackE):

> Environmental tracking dependent on the intersection of the environment's variability—which aspects of the environment vary, how (e.g., temporally each year, spatially at x scale) and how much—and a species' response to the environment via its proximate cues. If the varying components of the environment are not in the organism's set of cues, then the species may not 'track' per this definition. Environmental tracking at the individual-level is a purely plastic response to environmental variation and change; at the population-level tracking may also incorporate evolutionary change in the cue system, depending on both the timescales of study and the species' generation time. itsnotevoGiven our focus on responses to climate change, we consider environmental tracking here as a mainly plastic response (Bonamour *et al.*, 2019), though over longer timescales or in certain systems it should be shaped by selection (Franks & Hoffmann, 2012).

We have avoided yadstick (ii) purposefully and attempt to address that in this section as well, when we write (line r4yardstickS-line r4yardstickE):

> This is a foundational concept of the trophic mismatch literature (Visser & Gienapp, 2019), which often assumes the peak timing of a resource defines the ideal timing for phenological events dependent on that resource (e.g. egg laying dates dependent on caterpillar abundance, Visser & Both, 2005). For most phenological events, however, fitness outcomes are likely dependent on a suite of interacting forces—for example, egg laying dates for migratory birds may depend both on the timing of peak caterpillar abundance and the need to leave nesting grounds before winter.

This is a tricky topic and it's why we believe this paper could be useful to the field, but we appreciate we need to be exact and clear and we hope our updates to the text and Figure have

addressed this problem.

The section 1.3 on understanding variation in environmental tracking is rather long and doesnt offer up novel perspectives. I think it could be greatly reduced by briefly summarising some of the theoretical literature on the evolution of plasticity in response to environmental cues.

Agreed, we have overhauled this section and shortedned it considerably, see line plasS-line bhE.

I was surprised to see plasticity really only mentioned half way through the review (around line 245), given that along with any shifts in the environmental cues, this is the most important determinant of the phenological response at least in the short/medium term. I suggest that this could be mentioned earlier when you define environmental tracking. For instance, you could briefly outline the processes that can allow tracking, which I think are plasticity at the individual level, adaptation at the population level and species sorting at the community level. In lines 65-66 the mechanism underpinning a plastic response is defined and you might draw attention to that.

Agreed, we now mention plasticity much earlier (line mentionplastic) and here we discuss also the population level, "Environmental tracking at the individual-level is a purely plasticmentionplastic response to environmental variation and change; at the population-level tracking may also incorporate evolutionary change in the cue system, depending on both the timescales of study and the species' generation time." We have also overhauled the section on mentioned just above so that it opens with plasticity theory and focuses mainly on this.

I like the section on Tracking in Multi-Species Environments, as this is the first part of the ms that introduces some novel perspectives. I think the ms would be improved if the preceding components were edited down, so that you get to this point much sooner. In general I thought the second half of the ms was more stimulating and well-explained than the first.

Thanks ... we aimed to do that... XXX

Minor Comments
Environmental tracking: Where this idea is introduced (line 45) I think it might help to begin at the population level with a clear evolutionary definition of environmental tracking, where —B-b— is small following the equation in Chevin et al. 2010.

I feel like this could add confusion ... Megan, what do you think?

Line 6. What proportion? The cited paper by Cook et al. is just about phenology so doesn't support the general point. A paper by Amano et al. 2014 finds that UK plant species that shift less in terms of phenology have a greater tendency to range shift. I think this finding has been replicated in other systems but cant remember the reference.

We now Amano *et al.* (2014) on line line r4misc. TO DO ... add some more response here, as there are not that many other papers.

Line 14. And evolutionary theory, particularly Chevin et al. 2010 PLOS Biol.

We now cite Chevin *et al.* (2010) on line r4misc1.

Line 15. I think the terminology in this sentence is confusimg. From an evolutionary biology perspective plasticity has a clear meaning (a change in genotypes phenotype in response to the environment), but here I think it is being used to more vaguely imply flexibility, and I think flexibility would be a less loaded term. Also note that tracking can involve evolution.

Agreed, we now say "phenotypic flexibility (Piersma & Drent, 2003)" on line r4misc2.

Line 29-36. I agree that climate change has greatly exacerbated the non-stationary aspect of climate, but looking at historical records it seems as though climate is often somewhat non- stationary.

Agreed, we discuss this in the Box on 'Environmental variability & change.'

Line 56. I think a more precise/mathematical definition of cue quality could be helpful, e.g., something based on the sum of squares between optimum and actual event timing (RMSE?). Also note that the literature on the evolution of plasticity uses the term cue reliability to refer to the correlation between the environment of development and the environment of selection.

Need to fit in cue reliability! Megan, please add if you can ...

Line 62. Do you simply mean that in different locations if the individuals have the same reaction norms but environment differs then the outcome will differ? This could be explained in clearer language. Also there is a large literature by the likes of Scheiner, Lande, Chevin, Tufto, Hadfield on the evolution of cues and plasticity that goes uncited here.

In streamlining the manuscript we have deleted this sentence.

Line 64-67. Here the definition of tracking seems to be at odds with the evolutionary literature. The mechanism described is a plastic response to a cue, whereas in evolutionary biology tracking is usually with respect to a fitness optimum. This also seems to be at odds with your definition of fundamental tracking (line 48-49).

We have worked on this ... see above and see updated section line r1ass1-line moretrackE

Line 67. The organism is only expected to track the optimum proportional to the correlation between the environment of development and environment of selection.

In streamlining the manuscript we have deleted this sentence.

Line 84. Here you outline a series of papers that present information on phenological responses to temperature. However there is an absence of information on what the fundamental tracking or shifts in the optimum are doing. I think various methods exist for generating a yardstick (Visser and Both 2005) for fundamental tracking. One option is to use the response of resources. Alternatively, the estimation of the environmental sensitivity of selection (Chevin 2010) and use of this in prediction is an informative avenue (Vedder et al. 2013, Gienapp et al. 2013). We also use a space for time approach to estimate tracking of the optimum in plants (Tansey et al.

2017). In terms of environmental tracking another interesting perspective is that presented in the Amano et al. paper I mention above.

We agree this was not clear ... hopefully it is now. XXXX

Line 90-92. With respect to consumers tracking prey is this just the phenological shift shown? Here I think there is an opportunity to quantify whether tracking is adaptive (based on Ghalambour et als 2007 definitions of adaptive plasticity).

We have tried to address this in the section on 'Defining tracking' ... XXX ...

Line 174. See also Reed, T. E., Jenouvrier, S., & Visser, M. E. (2013). Phenological mismatch strongly affects individual fitness but not population demography in a woodland passerine. Journal of Animal Ecology, 82(1), 131-144.

XXX

Line 201. See Chevin et al. 2015.

XXX

Line 250. Is there a theory reference for this? I would have thought that the plastic response to each multivariate cue would be lower than the response to a single reliable cue.

XXX

Line 253. Evidence that the most plastic species have fared best Willis et al.

XXX

Line 420. This recommendation is a bit vague. Is there something quantitative that researchers should do?

XXX

Box. 2. An additional challenge for observational studies is teasing apart the influence of photoperiod. This may only be possible for spatiotemporal or experimental studies.

XXX

References

Amano, T., Freckleton, R.P., Queenborough, S.A., Doxford, S.W., Smithers, R.J., Sparks, T.H. & Sutherland, W.J. (2014). Links between plant species' spatial and temporal responses to a warming climate. *Proceedings of the Royal Society B-Biological Sciences*, 281.

Bonamour, S., Chevin, L.M., Charmantier, A. & Teplitsky, C. (2019). Phenotypic plasticity in response to climate change: the importance of cue variation. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 374.

Chevin, L.M. & Lande, R. (2015). Evolution of environmental cues for phenotypic plasticity. *Evolution*, 69, 2767–2775.

Chevin, L.M., Lande, R. & Mace, G.M. (2010). Adaptation, plasticity, and extinction in a changing environment: Towards a predictive theory. *Plos Biology*, 8.

Franks, S.J. & Hoffmann, A.A. (2012). *Genetics of Climate Change Adaptation*, vol. 46 of *Annual Review of Genetics*, pp. 185–208.

Menzel, A., Von Vopelius, J., Estrella, N., Schleip, C. & Dose, V. (2006). Farmers' annual activities are not tracking the speed of climate change. *Climate Research*, 32, 201–207.

Parmesan, C. (2006). *Ecological and evolutionary responses to recent climate change*, vol. 37 of *Annual Review of Ecology Evolution and Systematics*, pp. 637–669.

Piersma, T. & Drent, J. (2003). Phenotypic flexibility and the evolution of organismal design. *Trends in Ecology & Evolution*, 18, 228–233.

Simons, A.M. (2011). Modes of response to environmental change and the elusive empirical evidence for bet hedging. *Proceedings of the Royal Society B-Biological Sciences*, 278, 1601–1609.

Visser, M.E. & Both, C. (2005). Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B-Biological Sciences*, 272, 2561–2569.

Visser, M.E. & Gienapp, P. (2019). Evolutionary and demographic consequences of phenological mismatches. *Nature Ecology & Evolution*, 3, 879–885.